

# Psychological Review

**EDITED BY**

CARROLL C. PRATT  
PRINCETON UNIVERSITY

## CONTENTS

*The Appraisal of Child Personality: FLORENCE L. GOODENOUGH* ..... 123

*The Similarity Paradox in Human Learning: A Resolution:*  
CHARLES E. OSGOOD 132

*There is More than One Kind of Learning: EDWARD C. TOLMAN* ..... 144

*The Gestalt Theory of Expression: RUDOLF ARNHEIM* ..... 156

*'Superstitious' Behavior in Animals: W. N. KELLOGG* ..... 172

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879.

## COMPLETE MEMBERSHIP LIST

*Approximately 7,000 names*

The American Psychological Association maintains an address list of its members and affiliates, which is for sale providing the nature of its use is in conformity with the purposes of the Association.

Envelopes addressed ..... \$30.00  
(advertiser furnishes envelopes and pays express charges)

Addresses on tape, not gummed ..... \$15.00  
(suitable for a mailing machine)

### STATE LISTS

Priced according to number of names wanted

### SUPPLEMENTARY LISTS

*Approximately 3,000 names in total list  
Individual journal lists vary from 400 to 1,500 names*

The Association also maintains a list of subscribers who are not members of the Association (universities, libraries, industrial laboratories, hospitals, other types of institutions, and individual subscribers). The general list for all journals includes all types. Each single journal has a more specialized circulation.

For any one journal, envelopes addressed .... \$15.00

For any one journal, addresses on tape ..... \$10.00

For all journals, envelopes addressed ..... \$20.00

For all journals, addresses on tape ..... \$15.00

*For further information, write to*

**American Psychological Association**

1515 Massachusetts Avenue Northwest

Washington 5, D. C.

## THE PSYCHOLOGICAL REVIEW

THE APPRAISAL OF CHILD PERSONALITY<sup>1</sup>

BY FLORENCE L. GOODENOUGH

Almost half a century has passed since the publication in 1902 of Binet's *Experimental Study of Intelligence* (1). Had the scientific world then been ready to assess this work at its true worth, or had the need for a quick and practically feasible method of identifying those children whose mental capacity was so limited as to make it highly unlikely that they would profit greatly by the methods of school instruction designed for the generality been less pressing, the progress of mental testing might have followed a very different pattern from that with which we are familiar. But when, in 1904, Binet was assigned the task of developing an objective method for the selection of mentally retarded children to be placed in special classes, his attention was necessarily diverted from the broader study of qualitative differences in mental traits to the immediate need of a roughly quantitative device for the practical classification of children into ability groups. The tremendous enthusiasm engendered by Binet's scale for the measurement of intelligence after it had been brought to this country set the pattern for further work along the same line after Binet's untimely death in 1911. A not unnat-

ural result of this enthusiasm was the attempt to adapt the method which had proved so successful in the classification of subjects on the basis of scholastic aptitude and in the detection of the socially incompetent to the identification and measurement of other types of aptitude. The transition from the measurement of ability to the prediction of conduct appeared at first to be an easy one. It was assumed that by the use of procedures essentially similar to those employed for the measurement of intelligence or of scholastic accomplishment, equally successful devices might be developed for the classification of individuals into 'personality groupings' from which their social and emotional characteristics might be judged and their conduct in specific situations might be predicted with a satisfactory degree of probability.

On the basis of this hypothesis a wide variety of procedures for the study of personality differences were worked out. These methods are too well known to require more than brief mention. They include direct observations with or without the use of such predefined methods of recording as time-sampling or episode-sampling; standardized rating scales for use either by the subject himself, by his associates, or by teachers, parents or others in positions of responsibility; questionnaires of various kinds

<sup>1</sup> Address of the retiring president of the Division on Childhood and Adolescence of the American Psychological Association. Delivered at the fifty-sixth annual meeting of the Association, Boston, September 7, 1948.

including interest and attitude scales, check lists and preference ratings; records of honors and achievements and of school and court offenses; 'performance tests' such as those used by Hartshorne and May; ethical judgment tests, choice tests, and tests of information concerning various aspects of personal-social behavior. Most popular of all has been the so-called 'personality inventory' of which a scathing criticism was recently published by Ellis.

In all these methods certain implicit assumptions are involved. In the first place they are sampling methods. Now the essential features of a good sample are that it shall be adequate and unbiased. To be adequate, it is necessary that it represent all the essential features of the universe for which it stands. To be unbiased, these must be shown in their proper relationships and proportions.

It is at once apparent that the measurement and appraisal of the social and emotional aspects of behavior that we are accustomed to subsume under the heading of 'personality' present a number of special difficulties that are less often encountered in the measurement of ability. In the latter case, a direct approach to the task to be performed, with straightforward instructions, can be employed. We say to the subject, "I want to see how well you can remember numbers. Listen carefully and repeat these after me, just as I say them." But equally successful results are unlikely to be obtained if a similar method is followed in the study of an emotional reaction such as, for example, fear. Only a very silly experimenter would hope to learn much by saying to his subjects, "I want to see how frightened you can become. I am going to try to terrify you and I want you to become just as scared as you can possibly be."

Although this example is manifestly absurd, it illustrates a number of the

problems that have arisen from the unwise attempt to adapt the methods that have been found suitable for the study of abilities to the apparently similar but actually very different problem of the measurement of conduct where ability operates only as a limiting force. Because we cannot, in most cases, specify the situation when taking our test-sample, the universe of which the sample is presumed to be a representative part remains indeterminate. In the measurement of ability, the subject is directly motivated to do his best along some clearly specified line. The universe of which the test is assumed to be a sample is thus defined by the situation under which the test is given. In the case of an arithmetic test, for example, the subject's score supposedly represents his level of proficiency in the arithmetical skills covered by the tests *when he is motivated to do his best*. There is no assumption that he will always or even usually make a similar effort. In like manner his standing on an intelligence test is taken as an indication of the subject's *capacity* to perform difficult mental tasks, but there is no guarantee that his ordinary conduct will be on a par with his capabilities. The universe which the ability test attempts to sample is the *maximal* performance of which the individual is capable at the time of the test. A test of ability is, by definition, a test of the *limits of ability*.

Contrast this with the situation maintained when the sampling method is used for the measurement of those conduct-tendencies commonly known as 'personality traits.' What is the universe of which the sample is presumed to be a representative part?

In order to define a behavioral universe two things must be known; first, the kind of stimuli eliciting the responses in question, and secondly the level of response of which the behavior in question is assumed to be a sample.

I have specified the kind of stimuli rather than the type of responses because the former are under experimental control while the latter are not. The general experimental situation is, of course, regarded as part of the stimulus and must be specified along with it.

Unlike abilities which are defined in terms of the maximum, conduct-tendencies are commonly thought of in terms of the usual or typical. Now the determination of the maximum is affected only by errors of measurement but the determination of the mean requires a knowledge of the spread of the scores and thus calls for a relatively large number of samples, even when the error of measurement of the individual scores is small. As a rule when the subject is truly motivated to do his best, the experimenter can feel reasonably certain that the score earned on a particular task for which the error of measurement is not large will correspond fairly closely to the upper limit of which he is capable under the experimental conditions in question, even though only a single trial is given. But he will be on much less certain grounds if he assumes that the response to a single stimulation given *without* special incentive is representative of his subject's typical or usual response under roughly similar conditions. Unless the level of motivation is known, the level of response cannot be interpreted. One of the practical objections to the use of the sampling method when the aim is to determine the mean of a series of performances arises from the relatively large number of trials needed for a satisfactory approximation to the facts. While it is obviously impossible to ascertain the *absolute* maximum of a subject's performance, a reasonably satisfactory approximation to this figure can be had from a much smaller number of trials than is needed to establish his mean performance.

The choice of stimuli to be used in

the measurement of a complex performance will depend upon the experimenter's concept of the universe to be measured. Even when this universe is defined as an ability, rather than as a conduct-tendency, there will rarely be complete agreement among equally competent judges as to the nature and organization of that ability. The correlation between two tests of reading ability, for example, is rarely perfect, even after correction has been made for errors of measurement. Authorities differ in their concepts of the nature of intelligence, and the tests that they devise for its measurement show a corresponding lack of agreement in the results obtained by their use. But these differences, although significant, are small in comparison with those usually found among different psychologists in respect to their concepts of 'personality' and its organization, and in the magnitude of the correlation between tests of personality traits that bear the same name. Again the problem is one of sampling. Either the universes from which the samples were drawn, that is, the character of the traits which the different tests were designed to sample, are not the same in spite of their similar names, or the different test-samples were unequally biased. Whatever the explanation may be, the practical result is the same. The interpretation of an individual score in terms of a larger context becomes extremely hazardous.

In practical life, measurement usually implies some degree of appraisal. Five pounds of steak are worth more than a single pound. An all-wool blanket costs more than one that is but fifty per cent wool and the latter is more expensive than one that is only five per cent wool. In measures of ability a similar system of appraisal is commonly applied. A high score on an intelligence test is deemed preferable to a low score, and most people do not challenge the idea

that a direct linear relationship exists between social desirability and intelligence level. The same rule applies to most, if not all other measures of ability. The higher the score the more desirable it is thought to be, both from the standpoint of the individual and of his potential value to society. But what about such traits as dominance-submission or elation-depression or introversion-extroversion? At what point on the continuum does the ideal lie? Certainly not at either extreme. Nor is it likely that the relation between standing on these traits and social desirability can be assumed to follow a linear pattern. And finally it should be noted that appraisal of these traits cannot justifiably be made independently of other characteristics of the individual. The optimal level of social dominance is not the same for the imbecile as for the genius. The research chemist may lean strongly toward introversion but the traffic policeman must keep his attention on his surroundings. Superficially, at least, it would appear that although abilities may be measured and appraised independently of each other without giving rise to gross misunderstanding, conduct-tendencies such as those just mentioned cannot be properly evaluated except in relation to the total personality and to the context in which it operates.

For all these reasons, the sampling method of measuring and appraising conduct-tendencies has generally proved less successful than had been hoped. Correlations between tests designed to measure the same types of behavior have rarely been high enough to justify more than very coarse distinctions among individuals, and the relation between test scores and behavior outside the test situation has in most cases proved to be disappointingly low. Even when records of actual behavior such as are obtained by time-sampling or episode sampling are substituted for the formal test, the

difficulties do not disappear. Because there is no control and consequently no clear description of the conditions under which such records are obtained, dependable norms for general reference are very difficult to obtain. For example, Murphy found that the ratio between the number of episodes of sympathetic and unsympathetic behavior was more than four times as great in one of the two nursery-schools which she observed as in the other. This difference appeared to depend upon the total situation in which the children were placed and upon the character of the group atmosphere, rather than upon temperamental differences among the children making up the groups, although the latter may also have existed. Only ten per cent of the children in the first group equalled or exceeded the median number of unsympathetic responses displayed by those in the second group.

Not only do children behave differently, on the average, when the situation is changed, but such changes do not affect all cases in the same manner. An unpublished study carried out some years ago at the Institute of Child Welfare of the University of Minnesota by McConnon showed only very moderate correlations between the indexes of talkativeness obtained for nursery-school children in different situations. An outstanding example was that of a little girl who scarcely uttered a word during her three years' attendance at nursery-school in spite of continued efforts on the part of her teachers to stimulate conversation, but who proved to be quite a chatter-box at home.

Increasing realization of these and other difficulties likely to be encountered when the sampling method is employed for the study of personality differences has led to a search for other devices in which signs, rather than samples, are used as a basis for the interpretation and prediction of behavior. The basic

assumption underlying all such methods is that because of differing innate tendencies that have become still further differentiated as a result of modification through differing experiences, no two persons will observe the world from exactly the same angle. Because their perceptions differ, their responses to these perceptions will also differ. The behavior of every person thus provides us with a series of signs which, if properly interpreted, will enable us to understand much of the thoughts and feelings, the hidden motives and unverbalized meanings which Frank (3) has so cogently designated as the 'private world' of the individual.

The projective method may be described as a system of diagnosis that is based upon signs rather than upon samples. The usual procedure consists in presenting the subject with some kind of standardized material such as paints or other art-material, toys, pictures, inkblots and the like, with instructions that indicate only in a very general way what is to be done with the material, avoiding all suggestions as to the manner of dealing with it. Children are usually told simply that they may play with it in whatever way they like.

The growth of interest in projective methods since 1940 has been little short of phenomenal, paralleled only by the upsurge of interest in intelligence testing that took place between 1910 and 1920. Among the first 200 names whose fields of interest are specified in the 1948 Directory of the American Psychological Association (6), 23.5 per cent specifically mention projective techniques as one of their research interests, while among those in this group whose work is mainly with children, 37 per cent indicate such an interest.

With the kinds of methods used in the projective approach, most of you are entirely familiar. There are such formally developed tests as the Rors-

schach or the TAT, as well as a host of more informal procedures involving play with dolls or toy furniture, constructive work with blocks or other material, artwork of various kinds, the psychodrama, and so on. Although it has been customary to class all these under the general head of 'projective methods' it should be noted that the system of signs used in the Rorschach differs in a very essential way from that employed in most of the others. The interpretation of a Rorschach record is not based upon superficial resemblance between the kind of response made and the meaning that is assigned to it. Responses are classified according to the particular part of the blot chosen for interpretation, that is, the *location* of the response; its *determinant* such as color, form, shading and the like; the *content* of the response which, however, is not interpreted on a literal basis; and its *originality* as determined by comparison with the usual responses of other subjects of similar age and sex. The significance of all these factors is established by statistical comparison of the relative frequency of their occurrence in the records of subjects known to differ in respect to certain characteristics of personality or conduct. While it is unquestionably true that many of the interpretative Rorschach signs reported in the literature have been derived from very flimsy evidence, this evidence cannot be classed under the head of 'sympathetic magic' which is so strongly suggested by many of the procedures in common use. Although we no longer assume, as did our Puritan forebears, that the child who breaks the head of the doll chosen to represent his father thereby inflicts a similar injury upon the father himself, the assumption that a wish to do so as symbolized by the child's action is accepted without question by many. In like manner, the child who readily accepts an examiner's suggestion that he

break a set of fragile objects such as rubber balloons is said to be more 'aggressive' than another who prefers to keep the toys intact; one who likes to make mud-pies or who smears himself thoroughly with cold-cream is thereby assumed to display anal eroticism or some other form of sexual perversion. Perhaps the most glaring examples of anthropomorphic identification of the sign with the thing signified are to be found in the interpretation of children's art products. Not only are these interpretations usually made on an extremely superficial level but the symbolism itself is likely to be that of the adult rather than that of the child. The three-year-old sees no connection between the Easter egg that he so happily depicts with the aid of his new paints and the birth of a baby brother, but the adult, with his greater biological knowledge, too often reads into the child's painting all sorts of "birth fantasies," "sibling jealousies," and other conflicts of a similar nature. The yellow blob that the nursery-school child designates as a "lion" is much more likely to have been inspired by the illustration in his favorite picture book or, perhaps, by the treasured recollection of a one-time visit to the zoo than by an urge to give outward expression to a repressed desire to make a cruel and malicious attack upon his neighbor.

One of the most curious features of much of the work in the field of projective methods is the extent to which the workers have been blind to the projection of their own personality characteristics and their own cherished wishes and beliefs upon their interpretations of the behavior of their subjects. *For projection is a tool that cuts both ways.* If it is true that the person under observation projects his own inner feelings and attitudes upon the situation to which he responds, it is equally true that the person who observes him does the same.

Not only in his interpretations of the child's behavior, but in his arrangement of the situation and in the suggestions made in the course of the experiment can the examiner's bias be noted. When the only toy furniture provided in an experiment involving doll play consists of a bed, a toilet, and an arm-chair, it is not surprising to find that much of the play involves the use of the first two. When children are repeatedly told that they may break the toys if they wish, not much of importance can be assumed if they act upon the suggestion.

In pointing out these hazards of the projective approach I do not wish to give the impression that the method in general is a worthless one or that all the studies in which it has been employed should be thrown into the discard. On the contrary, it is just because I believe that the basic concepts upon which these procedures are based hold such great promise for the measurement and appraisal of conduct-tendencies as opposed to abilities that may or may not be given expression in actual behavior that I have tried to point out some of the false trails along which many are thoughtlessly treading. The attractiveness of these pathways, the apparent ease with which they may be traversed, their freedom from annoying statistics or other laborious procedures, and the substitution therefor of fascinating little anecdotes that are offered as proof of the very hypotheses on the basis of which they were chosen is certainly alluring. But science demands evidence of a different kind. Plausibility is not enough; for an anecdote which accords with a preconceived hypothesis may be selected in complete disregard of other evidence that runs contrary to it. And it is unfortunately true that many of the studies in this area have provided only anecdotal information to support the conclusions drawn. It is this disregard of the accepted rules of

scientific verification that has caused so many of the more 'hard-boiled' experimentalists to look upon the projective approach with scepticism. Although they may be willing to accept the idea in principle, they rightly disapprove of many of the methods comonly employed.

It should be unnecessary to state that the criticisms just made do not apply to all the work that is being done in this field. Recently, indeed, there have been a number of very promising attempts to improve and verify the procedures used as opposed to the more common attempts to rise at once to the heights of personality diagnosis upon the wings of fancy. The studies of factors affecting the character of doll play that have recently been carried out at the Child Welfare Research Station of the University of Iowa are examples. But much more work of this kind is needed if projective methods are to yield the information of which they are potentially capable.

At the beginning of this paper I referred to the early study by Binet of the personality characteristics of his two daughters. Most of you are doubtless familiar with this study by hearsay at least, but its importance is such that you will forgive my recalling it to your attention. In my opinion it remains the best as it almost certainly is the first example of the use of the projective approach to the study of personality differences. As some of you will recall, Binet was interested in demonstrating that elaborate procedures and complex instruments are not always necessary for the attainment of results that are scientifically valid. Chiefly, however, he wished to show that the essential mental characteristics of an individual are betrayed in his ordinary responses to the small affairs of everyday life as well as in the more precise experiments of the formal laboratory. But Binet did not make the mistake of assuming that these

subtle indications of personality characteristics are simple mirror images of the behavior tendencies of which they are the overt signs, nor did he rely on unaided intuition for the identification of these signs or for their interpretation. Always he had the wisdom to *let facts lead*, and the humility to accept his data at their face value without attempting to force his figures into some preconceived mold of his own making. Neither did he attempt to draw far-reaching conclusions from half-a-dozen hastily noted facts. Instead he tried experiment after experiment, checked and re-checked his data until the internal evidence of his results was so striking that no one can question it.

I have dwelt upon this study of Binet's at some length, because, as I stated before, the work of the quarter-century just passed has pretty well convinced me that the sources of human conduct lie too far below the surface to be easily reached and sampled by the direct methods that we have been accustomed to employ in the measurement of abilities. Nevertheless, these hidden sources are not wholly inaccessible. They reveal themselves in signs, the meaning of which is not immediately apparent but which can be learned by patient study. For we must rid ourselves of the mistaken notion that these signs take the form of open symbolism that can be interpreted on the basis of its superficial resemblance to some designated form of behavior or type of personality. Let me illustrate by citing the results obtained by Binet for a single one of the twenty or more experiments carried out with his two daughters. The experiment consisted in asking the girls to write twenty words, just as they came to mind and as rapidly as possible. Immediately after the words had been written, the girls were questioned individually as to the particular associations which brought each word to mind. Their explanations

were recorded without comment. After the lapse of several days the experiment was repeated until sixteen lists, totalling 320 words for each subject, had been secured. Binet then examined the lists with care to note what aspects of the words and of the explanations given for their choice appeared to be significant. Here, as throughout the series of experiments, it is noteworthy that he made no *a priori* assumptions. He permitted the data to speak for themselves.

On the basis of his observations, six categories were selected as probably meaningful. These categories, and the number of words in each as given by the two girls are as follows:

1. *Unexplained words*, that is, words for the choice of which the subjects were unable to give any account whatever. They seemed to come by themselves, without other associations. In the record of the older daughter, Marguerite, there were only 15 words of this class as opposed to 84 for Armande, the younger daughter. Binet notes, among other things, that although separate trials with dictated material showed that the usual speed of writing of the two girls was approximately equal, during these experiments Armande always wrote more rapidly than Marguerite. Through his questioning Binet was able to show that this difference was largely attributable to what we may call "mental set." Marguerite's attention was, for the most part, directed to the sense of what she was writing. Armande tended to think of words as words, with less attention to their meaning.

2. *Names of objects immediately present to the senses*. In Marguerite's list, words of this class account for 120 of the total, in Armande's list for only 30.

3. *Words referring to the self*. Under this head, Binet included only those words that had to do with the subject's

physical appearance or with her own clothing. Marguerite wrote 15 words to which self-reference was imputed; Armande, none.

4. *Memories*. These were words definitely stated by the girls to have reference to some past experience. Of these, Marguerite's list included 172, Armande's but 88.

5. *Abstract words*. Armande wrote 70 words of this type, Marguerite only 12.

6. *Words with imaginary reference*. As used by Binet, this category included only those cases in which a 'fictive image' of some object or scene appeared to rise spontaneously in the mind of the subject and suggest the word that was written. Memories of actual occurrences were not included. No such images were reported by Marguerite, but Armande's total was 23 unquestionable cases as well as a number of other instances in which the presence of the image was less certain.

The differences in the mental processes of the two girls as revealed by these figures is at once so striking and so internally consistent that many of us might feel that nothing more was needed for an elaborate personality diagnosis. Not so Binet. He continued his experiments, using as tasks the writing of sentences and of stories, the description of objects, reports of remembered events, measures of reaction-time, the interpretation of pictures, tests of sentence completion, the identification of errors in printing, and many others. Always, however, Binet's method of treating his results was determined by the study of what actually occurred and not by any preconceived notion of what ought to occur. He did not manufacture signs in advance, but he searched diligently for those that were there. By checking these signs, one against another, he was

eventually led to conclusions so inevitably *right* that they admit of no doubt.

All of us who are interested in the projective approach to the study of personality differences may well look upon these early experiments by the father of mental testing as models of scientific exactness. Too often we have been lured by the hope that surface appearances will reveal the hidden depths that lie beneath. We have spoken glibly of 'private worlds' but have ignorantly assumed that intuition alone can provide the key by which their privacy may be laid bare for public inspection. We have sought in the projective techniques for an easy road to wisdom.

To Binet we are indebted for our most important method of measuring intelligence. He has also demonstrated that the projective method is not necessarily a device for those who are either unwilling or incapable of undertaking the patient search for truth that is the mark of

the genuine scientist, but is susceptible to as complete and objective verification as the most rigorous critic can demand.

#### REFERENCES

1. BINET, A. *L'étude expérimentale de l'intelligence*. Paris: Alfred Costes, 1902. Pp. 307.
2. ELLIS, A. The validity of personality questionnaires. *Psychol. Bull.*, 1946, **43**, 385-440.
3. FRANK, L. K. Projective methods for the study of personality. *J. Psychol.*, 1939, **8**, 389-413.
4. HARTSHORNE, H., & MAY, M. A. *Studies in the nature of character. I. Studies in deceit*. New York: The Macmillan Co., 1928. Pp. xxii + 306.
5. MURPHY, LOIS B. *Social behavior and child personality: an exploratory study of some roots of sympathy*. New York: Columbia University Press, 1937. Pp. ix + 344.
6. WOLFLE, HELEN M. (Ed.). *American Psychological Association, 1948 Directory*. Washington, D. C.: American Psychological Association, 1948. Pp. viii + 429.

[MS. received September 28, 1948]

# THE SIMILARITY PARADOX IN HUMAN LEARNING: A RESOLUTION

BY CHARLES E. OSGOOD

*University of Connecticut*

Behavior is a continuous, fluid process, and activities learned in the laboratory are as much a part of it as a trip to the county fair. The segments which an experimenter arbitrarily selects for analysis are inextricably imbedded in this expanding matrix and are interpretable only in terms of its interactions. Transfer and retroaction experiments are explicit attempts to gauge these interactions, and the similarity variable—that is, the homogeneities existing among the materials successively practiced—turns out to be the most important factor as well as the most puzzling.

The classic statement of the relation between similarity and interference in human learning, as found in most textbooks in psychology, is that "the greater the similarity, the greater the interference." Although this law is traceable mainly to the work of McGeoch and his associates (10, 13, 14), there are many other experiments which superficially appear to substantiate it. When carried to its logical conclusions, however, this law leads to an impossible state of affairs. The highest degree of similarity of both stimulus and response in the materials successively practiced is that where any simple habit or S-R association is learned. The stimulus situation can never be precisely identical from trial to trial, nor can the response, but they are maximally similar—and here the greatest facilitation (ordinary learning) is obtained. *Ordinary learning, then, is at once the theoretical condition for maximal interference but obviously the practical condition for maximal facilitation.* Here is the fundamental paradox, and this paper suggests a resolution.

## EMPIRICAL LAWS OF TRANSFER AND RETROACTION AS FUNCTIONS OF SIMILARITY

Transfer and retroaction in human learning are among the most extensively cultivated fields in experimental psychology, yet there are no clear-cut generalizations which satisfactorily bind the data together. The difficulty may be traced in part to the bewildering variety of procedures, materials and experimental designs employed by different investigators, a phenomenon perhaps characteristic of a young science. But some of the confusion can also be laid to the fact that in a large proportion of experiments the theoretically relevant relations are patently unspecifiable: the subjects merely learn List A and then List B, or Maze I and then Maze II, and either positive or negative effects may result, depending upon quite unanalyzable conditions. The purpose of this paper is to clarify the similarity function in human learning, and to accomplish this end only those experiments can be utilized wherein the *locus* of similarities is specifiable, as being between stimulus members, response members, or both. This analytic approach, although it may be considered inappropriate by some theorists and makes use of only part of the data, does give rise to a coherent and consistent picture.

When *transfer* is studied, one is interested in the effect of a specifiable prior activity upon the learning of a given test activity. When *retroaction* is studied, one is interested in the effect of a specifiable interpolated activity upon the retention of a previously learned ac-

tivity. In both cases the experimenter arbitrarily "lifts" segments of a continuing process for analysis, and it would be expected that common laws would apply to both samplings. In the present context it can be shown that identical functions of similarity apply to both transfer and retroaction data, which simplifies the theoretical task considerably. Figure 1 gives symbolic representation to three basic learning paradigms, *A* that in which stimulus members are varied in the materials successively practiced while responses are functionally identical, *B* that in which responses are varied and stimuli are functionally identical, and *C* that in which both stimulus and response members are simultaneously varied. It will be seen that in so far as similarity relations are concerned, the test for transfer is simultaneously the interpolated activity when the entire retroactive sequence is followed. The term "functional identity" is used here to make explicit the fact that *true* identity among either stimulus or response processes is a will-o-the-wisp, approached but never attained. Functional identity of stimuli in successive trials or tasks exists when the situation is objectively constant (*i.e.*, when the same stimulus nonsense syllable appears on the screen or the same choice point is approached on repeated trials in the maze); functionally identical responses are those which the experimenter, at any given level of analysis, scores as being the same (*i.e.*, no matter how the sub-

ject says CYF or how the rat maneuvers about a turn, it is scored 'correct'). Functional identity thus becomes the limiting case of maximum similarity.

1. Let us first consider *paradigm A*, the condition in which stimulus similarity is variable and responses are functionally identical. The transfer portion of this paradigm will be recognized as nothing other than a symbolic statement of *stimulus generalization*. In Hovland's classic study (9), for example, a galvanic skin response is first conditioned to a tone of a certain frequency ( $S_1-R_1$ ), then the test tone is presented and the extent to which the same response is made to it measured ( $S_2-R_1$ ). Hovland found that the greater the similarity between practice and test stimuli, the greater the amount of generalization (or positive transfer). The same results are regularly found wherever this paradigm can be identified, whether the materials be motor or verbal, meaningful or nonsense, or of any other nature. McKinney (15) required subjects to respond with a correct letter upon seeing each of four geometrical designs and then measured transfer of the same responses to alterations of these designs; when Yum (25) varied the similarity of visually presented nonsense-syllable stimuli, positive transfer was the result, the magnitude increasing with stimulus similarity.

While retroaction data derived from this paradigm are not so extensive, the available evidence is consistent in revealing *facilitation*. Hamilton's (7) subjects learned lists of paired-associates in which the stimuli were geometrical forms and the responses were nonsense syllables. Although responses were 'identical' on original and interpolated lists, the stimulus forms varied from 'identity' through two degrees of similarity, as independently indexed in terms of generalization, to complete neutrality. The magnitude of retroactive facilitation

#### Transfer and Retroaction Paradigms

##### TRANSFER

Paradigm A	$S_1 \rightarrow R_1$	$\underline{S_2} \rightarrow R_1$	$S_1 \rightarrow R_1$
Paradigm B	$S_1 \rightarrow R_1$	$S_1 \rightarrow \underline{R_2}$	$S_1 \rightarrow R_1$
Paradigm C	$S_1 \rightarrow R_1$	$\underline{S_2} \rightarrow \underline{R_2}$	$S_1 \rightarrow R_1$

##### RETROACTION

FIG. 1. Paradigms indicating the locus of variation among the successively practiced materials. *A*, stimulus variation; *B*, response variation; *C*, simultaneous stimulus and response variation.

decreased regularly as similarity among the stimulus members decreased, effects of approximately zero magnitude being obtained with neutral stimuli. The empirical law for this paradigm is: *where stimuli are varied and responses are functionally identical, positive transfer and retroactive facilitation are obtained, the magnitude of both increasing as the similarity among the stimulus members increases.*

2. The situation in which stimuli are constant and responses are varied, *paradigm B*, is the standard associative and reproductive inhibition paradigm and, as might be expected, a large number of experiments (*cf.* 1, 6, 21) testify to the fact that *interference* is produced under these conditions. However, there is also a large body of evidence showing positive transfer under the same conditions. The latter evidence may be discounted on two grounds: (a) In many cases the so-called transfer response has been *learned previous* to the experimental situation. In many of Tolman's sign-learning studies, for example, animals trained to traverse the route to a goal by one path or means, such as running, will shift readily to another means, such as swimming, if the original behavior is blocked. Similarly, Wickens (23) has shown that a human subject who has learned to avoid the shock which follows a tone by an extensor movement of his finger, when his palm is down, 'transfers' immediately to a flexion movement when his hand is then placed palm up. In such cases, the new learning in the experimental situation is the sign-value or meaning of the distinctive cue. A variety of overt behaviors has previously been associated with this mediation process—the human subject brings to the experiment a rich repertoire of pain-avoiding movements, and he would lift his head without new training if his nose were inserted between the electrodes!

(b) In other cases what is measured as

positive transfer under conditions fitting this paradigm can be shown to be attributable to '*practice effects*,' *i.e.*, the subject is learning how to learn nonsense syllables or learning how to learn mazes, and these general skills or habits counteract the interference inherent in the design. Siipola (20), for example, obtained small amounts of positive transfer for a code-substitution task, yet concluded from the large numbers of intrusions that actual negative transfer was being masked by a general 'practice effect.'

Bugelski (2) required his subjects to learn an original list of 10 paired nonsense syllables (such as *toc-nem*) and then interpolated three additional lists, the experimental subjects having identical stimuli and varied responses (such as *toc-rul*) and the control subjects having both members varied (such as *cos-rul*). Although insignificant amounts of positive transfer to successive lists were obtained in both conditions, the inherent interfering character of the stimuli-identical paradigm was revealed in the fact that the experimental subjects showed a marked decrement upon relearning the first list while the controls showed continued facilitation. Clearest evidence for negative transfer and retroactive interference under the conditions of this paradigm is offered in a recent monograph by Underwood (21). In measuring transfer, subjects learned 0, 2, 4, or 6 lists of meaningful paired-associates *prior* to learning a test list; in measuring retroaction, 0, 2, 4 or 6 interpolated lists were learned *after* the original learning of the same test list; in both cases, recall of the test list was measured after a delay of 25 minutes. Both negative transfer and retroactive interference were found, increasing in magnitude with the number of prior or interpolated lists having the same stimulus members but different responses.

But what about the *degree* of simi-

larity among the varied responses in this paradigm? Perhaps because of the difficulty in defining response similarity, there are relatively few data here. In a recent experiment by Osgood (17), original learning of a set of paired letter-pairs and meaningful adjectives (such as *c.m.—elated*) was followed by three types of interpolated items, each subject serving as his own control by learning an equal number of items in each similarity relation (such as *c.m.—high, c.m.—left, or c.m.—low*); all subjects finally relearned the original list. Although interference was obtained under all conditions, it was significantly *less* for similar meaningful relations. One of the conditions of Bruce's (1) extensive investigation with nonsense-syllable paired-associates substantiates this finding. We may now state the empirical law for this paradigm: *where stimuli are functionally identical and responses are varied, negative transfer and retroactive interference are obtained, the magnitude of both decreasing as similarity between the responses increases.*

3. *Paradigm C*, where both stimuli and responses are simultaneously varied, is directly generated when the standard memory drum is used and lists of material are learned in constant serial order. Similarities are between items having the same serial position on successive lists, and each item serves simultaneously as a response to the preceding item and a stimulus for the succeeding item. Whatever interpolated lists are given, stimulus and response similarities must be simultaneously varied through the same degrees. McGeoch and McDonald (13) and Johnson (10) have employed this procedure with meaningful materials, finding retroactive interference to increase with the degree of similarity. Melton and Von Lackum (16) report the same result for nonsense syllables. McGeoch and McGeoch (14) and Johnson (10) find the same result

to hold for transfer when this paradigm is used.

An important experiment by Gibson (6) also fits this paradigm. Her materials and procedures were identical with those reported above for Hamilton (7). The Gibson experiment was actually the first of the series. Visual stimulus forms were varied through independently measured degrees of generalization, as was the case in Hamilton's study, but here responses were different and neutral. Negative transfer and retroactive interference were obtained, their magnitudes decreasing as stimulus similarity decreased and approximating zero with neutral stimuli. It should be noted that in both studies approximately zero transfer or retroaction was found when stimuli were neutral, regardless of response identity or difference. The empirical law for this paradigm: *when both stimulus and response members are simultaneously varied, negative transfer and retroactive interference are obtained, the magnitude of both increasing as the stimulus similarity increases.*

There are a considerable number of substantiating studies which have not been cited here, but if this writer's survey of the literature has been adequate, *there are no exceptions to the above empirical laws.* There are few studies where more than one relation is systematically explored, with the same materials, procedures and subjects, and for this reason it is difficult to quantify these relations. An exception is a study by Bruce (1). One set of nonsense pairs (such as *req-kiv*) was learned by all subjects and transfer to several variations was measured: where stimuli were varied and responses were constant (*zaf-kiv* or *reb-kiv*) positive transfer was found as compared with a control condition, the amount being greater when stimuli were more similar: where responses were varied and stimuli were constant (*req-vor*), negative transfer was found. The

condition in which stimuli were constant and responses were highly similar (*req-kib*) was slightly superior to the control condition (both members neutral). Although this result appears to contradict the empirical law for this paradigm, it will be found to fit the hypothesis presented in the latter part of this paper: if ordinary learning is to be theoretically feasible, high degrees of response similarity must yield facilitation.

#### ATTEMPTED INTEGRATIONS OF THE DATA

A series of attempts to integrate the facts of transfer and retroaction can be traced in the history of this problem. As early as 1919 Wylie (24) had made a distinction between stimulus and response activities, stating that the transfer effect is positive when an 'old' response is associated with a new stimulus but negative when an 'old' stimulus must be associated with a new response. "Old" in this context merely means that the member in question has previously been associated with another stimulus or response. This principle is valid, of course, within the limits of its gross differentiation. But (a) it takes account of neither stimulus nor response similarities and (b) it leaves the fundamental paradox untouched. Since successive responses are never precisely identical, even in ordinary learning, we are always associating stimuli with 'new' responses and hence should inevitably get negative transfer.

Robinson was one of the first to perceive clearly this paradox and in 1927 he offered what is now known as the Skaggs-Robinson Hypothesis as a resolution. As shown in Fig. 2, this hypothesis states that facilitation is greatest when successively practiced materials are identical (point *A*); facilitation is least, and hence interference maximal, with some moderate degree of similarity (point *B*); and facilitation increases again as we move toward neutrality.

#### THE SKAGGS-ROBINSON HYPOTHESIS

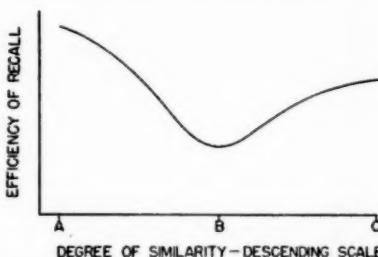


FIG. 2. The Skaggs-Robinson Hypothesis: point *A* specifies maximum similarity (identity) among the successively practiced materials; point *C* minimum similarity (neutrality); point *B* merely indicates the low point in the curve for efficiency of recall.

(point *C*) but never attains the original level. Note that while point *A* defines maximum similarity (identity) and point *C* defines minimum similarity (neutrality), point *B* actually specifies no degree of similarity at all, but merely says that somewhere there is a low point in the facilitation curve. Several experiments (3, 4, 8, 11, 19) combine to give rough validation to this poorly defined hypothesis, especially the *A-B* sector of it.

The series of studies by McGeoch and his associates (10, 13, 14) ran into direct conflict with this hypothesis and the experimental evidence supporting it. Using meaningful words, they consistently found that as the judged similarity of the original and interpolated materials increased, interference also increased. The highest degree of similarity they could obtain, where close synonyms appeared on the two serial lists, yielded the most interference. There was no evidence here of facilitation as one approached identity. In *The Psychology of Human Learning* (12) McGeoch offered two alternative rapprochements between his data and the Skaggs-Robinson Hypothesis: (1) He

distinguished 'similarity of meaning' and 'degrees of identity' as two different dimensions of similarity, each having a different interference function. This distinction was suggested by the fact that some of the experiments supporting the hypothesis (8, 11, 19) had employed numeral and letter combinations with similarity indexed by the number of identical elements. Unfortunately, in other substantiating studies, materials were used in which identical elements were no more readily specifiable than with meaningful words. Dreis (4), for example, used code-substitution, and Watson (22) used card-sorting. Furthermore, this type of resolution implies an analysis of meaningful similarity that would segregate it from identity of elements, and this has not been done. (2) At a later point, McGeoch tried to resolve the difficulty by stating that his results applied only to the portion of the Robinson Curve between *B* and *C*, *i.e.*, that the maximum similarity of his materials only reached point *B*. However, given the multidirectional shape of this theoretical function and the fact that point *B* defines no degree of similarity, not only could any obtained data be fitted to some portion of it, but it could always be argued that the similarity of one's materials fell *anywhere* between *A* and *C*. In other words, this second suggestion is incapable of either proof or disproof.

Perhaps the clearest experimental evidence against either of McGeoch's resolutions appears in the results of a recent experiment by the writer (17). *Also* using meaningful materials in the traditional retroaction paradigm, interference was found to *decrease* as the meaningful similarity among the response members increased. Not only would these results seem to fit 'degrees of identity' rather than 'similarity of meaning' as the functioning dimension, despite the nature of the materials used, but they fall within

the *A* to *B* sector of the theoretical curve.

Quite apart from the apparent negative evidence in the McGeoch studies, the Skaggs-Robinson Hypothesis is inadequate on several grounds. It does, to be sure, allow ordinary learning to occur. But (a) it contains a dual function of facilitation in relation to similarity without specifying at what degree of similarity the shift occurs; (b) no specification is made of the locus of similarities within the materials practiced (whether among stimulus members, response members or both), and we have seen that both the direction and the degree of either transfer or retroaction are empirically predictable from such specification. One of the most recent attempts to integrate these data has been made by Gibson (5). She followed Wylie's lead in differentiating between stimulus variation and response variation, and she added to this picture the refinement of stimulus generalization, derived from Pavlovian conditioning principles. Gibson's two theoretical laws were: (1) if responses are *identical* facilitation is obtained, its amount increasing with the degree of stimulus generalization (similarity); (2) if responses are *different* interference is obtained, its amount increasing with the degree of stimulus generalization (similarity). These hypotheses fit much of the data in the field and further serve to integrate the phenomena of human learning with those of the animal laboratory. But they are insufficient. (a) No account is given of the *degree* of response similarity, and this appears as one of the relevant variables. (b) We have one function (increasing interference) when responses are *different* and another (decreasing interference) when responses are *'identical'*; and one would anticipate, therefore, a strange, abrupt shift in function somewhere along the line as the degree of response difference is reduced.

(c) The fundamental paradox remains: responses can never be truly identical but must always be different to some degree, yet ordinary learning can occur.

#### THE TRANSFER AND RETROACTION SURFACE

The formulation proposed here makes full use of Gibson's analysis, but, utilizing data which have recently become available, goes beyond it. It is quite literally constructed from the empirical laws presented above, and this can be demonstrated by use of Fig. 3, which provides a rational framework within which the data can be integrated. The vertical dimension represents the direction and degree of *either* transfer or retroaction; degrees of response similarity are distributed along the horizontal dimension. The parenthetical numbers refer to the sequence of steps to be followed in allocating the data.

Let us first consider the ordinary learning of an association, the case in which the same materials are used for original and interpolated activities. Here functionally identical stimuli and

responses are successively repeated and maximal facilitation is obtained, allowing us to locate the first point as shown (number 1). The phenomena of positive transfer (stimulus generalization) and retroactive facilitation when responses are identical and stimuli varied are represented by the series of open circles (number 2): as the degree of stimulus similarity decreases from 'identity' less and less facilitation is obtained, effects of zero magnitude being found when stimuli are neutral. Data reported by Hovland (9) and Hamilton (7) are typical. As pointed out earlier, the fact that Hamilton and Gibson (6) used the same materials and procedures, with the single exception that responses were the same in the former case and different in the latter, provides an extremely useful comparison (see number 3); where stimulus members are neutral, effects of approximately zero magnitude are obtained in both experiments, allowing us to link the Gibson and Hamilton data together on the zero-effect base line. In other words, variations in the relation between response members are of no

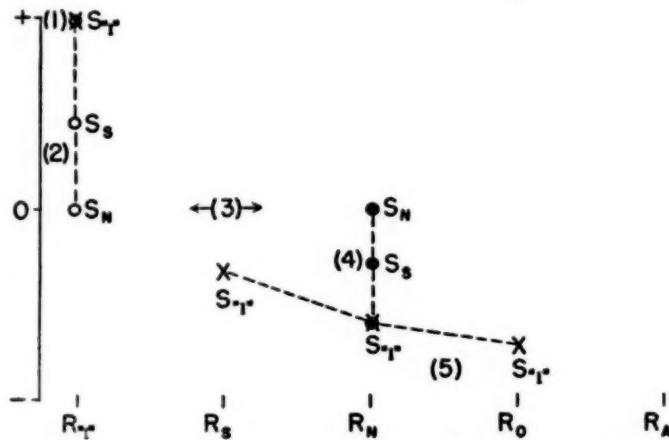


FIG. 3. Allocation of experimental data: vertical, direction and degree of either transfer or retroaction; horizontal, degrees of response similarity. Numbers in parentheses refer to step in analysis followed in text.

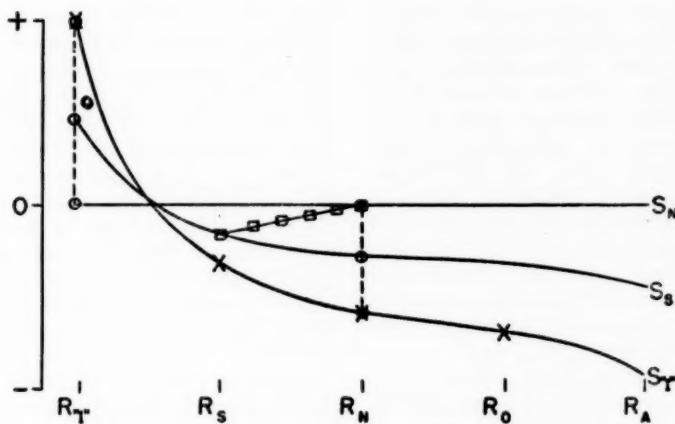


FIG. 4. Family of stimulus-relaton curves constructed from data in Fig. 3; series of open squares represents data obtained by McGeoch and his associates (see text).

consequence when stimulus members are completely unrelated. The Gibson experiment itself, along with other substantiating studies, provides data for the condition in which responses are different and neutral while stimulus similarity is varied. Here negative transfer and retroactive interference are regularly obtained, increasing in magnitude as the similarity of the stimulus members increases, and these data are represented by the series of solid circles (see number 4). There remains to be included the condition in which stimuli are constant and response similarity is varied. The fact that 'identity' of stimulus and variation of response yields negative transfer and retroactive interference is amply testified to by a number of studies (1, 6, 21). Experiments in which the *degree* of response similarity is systematically varied, as those by Bruce (1) and Osgood (17), show that interference is *less* for similar responses than for neutral ones. Since the latter study included a condition in which responses were neutral and stimuli functionally identical, thus matching the final point of the Gibson data, it is pos-

sible to link the two sets of facts together. These data are represented by the connected series of X's (number 5).

The pattern of empirical points established here sharply limits the possible theoretical functions that can be generated. By visually tracing the series of X's, for example, including the point for ordinary learning, a fairly well-defined curve becomes apparent, this curve representing the function for stimulus 'identity.' A family of such stimulus-relaton curves has been constructed to fit both these empirical points and the requirements of common sense, and they appear in Fig. 4. The function for *stimulus neutrality* is a straight line of zero effect, reflecting the reasonable fact that response variations are of no consequence when successive stimulus situations are completely unrelated. Given this as a zero-effect base line, increasing the similarity among stimuli yields a progressive maximization of *both* facilitation and interference, the actual direction of the effect being dependent upon response relations. The greatest facilitation and the greatest interference are possible only with functional *stimulus*

*identity.* Intermediate transfer and retroaction effects fall between these limits depending upon degrees of stimulus similarity. The points for antagonistic responses, showing a final, sharp increase in interference, are admittedly hypothetical. However, the writer has recently reported (18) evidence for a special form of *reciprocal inhibition* associated with the successive learning of meaningfully opposed responses. The assumption is made here that this inhibitory effect is maximal when responses are directly antagonistic.

But how do the classic findings of McGeoch and his associates fit this hypothesis? In a real sense, they serve as a crucial test of it, being both well substantiated and in apparent conflict with other results. It will be remembered that these investigators employed a method wherein the similarity of *both* stimuli and responses varied simultaneously and through the same degrees, ac-

tually from neutrality of both to high similarity (but not identity) of both. As may be seen from the row of open squares in Fig. 4, the present hypothesis *must* predict gradually increasing amounts of interference under these conditions, and this is precisely the result obtained in these studies.

Although Fig. 4 provides a useful method of demonstrating the congruence of empirical data and theoretical functions, it does not offer a clear picture of the hypothesis as a whole. To do so requires a three dimensional form, representing stimulus similarity, response similarity and degree of effect as simultaneously interrelated variables. Figure 5 presents what may be termed *the transfer and retroaction surface*. The vertical dimension represents the direction and degree of either transfer or retroaction, both having been shown to have identical functions of similarity; the width of the form represents stimulus similarity, from

### THE TRANSFER AND RETROACTION SURFACE

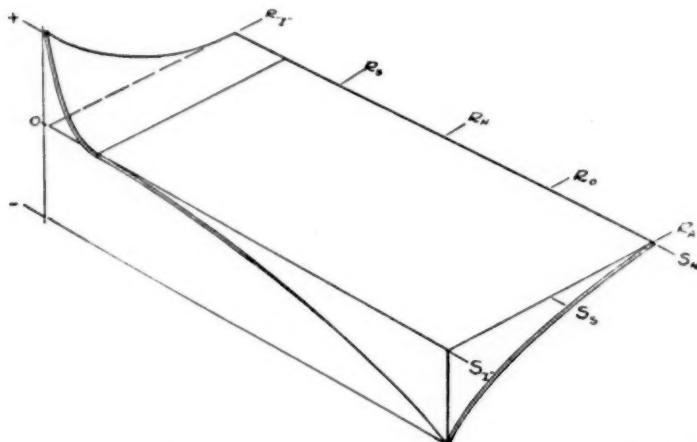


FIG. 5. The transfer and retroaction surface: medial plane represents effects of zero magnitude; response relations distributed along length of solid and stimulus relations along its width.

functional identity to neutrality; its length represents response similarity, varying from functional identity, through neutrality, to direct antagonism. The median horizontal plane indicates effects of zero magnitude, and it may be seen that the condition of stimulus neutrality is co-extensive with this plane regardless of response variations while the remainder of the surface intersects this plane at a point between response 'identity' and response similarity. Finally, it is apparent that we have here a smooth, unbroken sequence of transfer and retroaction functions, facilitative relations rising above the median plane and interfering relations falling below it. There are no reversals in these functions nor any abrupt shifts between identity and similarity. Identity becomes merely the limiting case of maximal similarity.

#### CERTAIN ADVANTAGES OF THIS HYPOTHESIS

By way of summary, certain advantages which this hypothesis offers in comparison with those which have preceded it may be indicated. (1) *All existing empirical data in the field are consistent with it and find representation upon the transfer and retroaction surface.* This statement is by necessity limited to those data wherein the locus of the similarities is specifiable and also by the adequacy of the writer's survey of the literature. The first limitation is not a serious one. If results can be shown to be lawful, and hence predictable, when such specification of the similarity relations is possible, the conflicting and confused results obtained under unspecifiable conditions are presumably attributable to unanalyzable variations in the paradigms employed. Witness the conclusive inconclusiveness on the question of formal discipline! This state of affairs illustrates why it is so difficult to make recommendations for efficient human learning in practical situations.

What, for example, are the loci of similarities when the student simultaneously studies French and Spanish?

(2) *The phenomena of both transfer and retroaction are integrated within a single framework, in so far as the similarity variable is concerned.* It is common textbook procedure to study transfer under learning and retroaction under forgetting, as if these processes were somehow different in kind. The present analysis, it is felt, is a step in the direction of integrating the problems of human learning. Another step in the same direction is also suggested here: distinctions are often made in terms of meaningful vs. nonsense materials, meaningful similarity vs. degrees of identity, and so on. It should be pointed out that data substantiating each of the three empirical laws derived above have been obtained with meaningful and nonsense materials, with materials varying in terms of meaningful similarity as well as degrees of identity. There is here, of course, the underlying problem of defining similarity. It may be defined operationally in terms of generalization (*cf.* Gibson, 5), although this definition is inherently circular since the phenomenon of generalization is nothing other than a case of positive transfer with functionally identical responses. Any precise behavioral definition of similarity will require much more knowledge of the nervous system than we have at present. In practice, degrees of similarity have been specified informally by experimenters or formally by a sample of judges, which probably suffices for our present rather gross purposes.

(3) *Although constructed directly from existing empirical evidence, this hypothesis does go considerably beyond it, predicting phenomena that have not as yet been observed.* For one thing that portion of the transfer and retroaction surface where increasing similarity of response (high degrees) is accom-

panied by increasing facilitation remains to be explored by standard procedures, the Robinson group of studies having used a memory span technique.<sup>1</sup> It will also be noticed that the theoretical surface requires that, regardless of the degree of stimulus similarity, all functions must become facilitative at precisely the same degree of response similarity, somewhere between identity and high similarity. In other words, just as the degree of response variation is inconsequential when stimulus members are neutral, so there must exist (according to this hypothesis) some definite degree of response similarity for which all variations among stimuli will yield zero effect. This is a novel but necessary prediction from theory that sets an intriguing experimental problem. It is not inconceivable that this common shift-over from facilitation to interference at a certain degree of variation among responses may reflect a basic characteristic of the nervous system,—but this is all assuming that the present hypothesis will be found valid in terms of constantly accruing facts.

(4) Finally, this hypothesis resolves the fundamental paradox with which this paper began—the fact of ordinary learning becomes theoretically feasible. The transfer and retroaction surface describes a system of curves within which the condition of ordinary learning, with functionally identical stimuli and responses in the materials successively practiced, is continuous with other relations. Identity is here merely the limiting case of maximal similarity, and no

<sup>1</sup> An as yet uncompleted investigation by Mark W. Harriman at Johns Hopkins University appears to be filling in this gap in our empirical knowledge. With functionally identical stimulus members, responses on original and interpolated lists are varied by extremely small degrees, such as having the singular and plural of the same word on two lists, and the predicted results seem to be forthcoming.

abrupt shifts of function are required to account for the fact that learning occurs.

#### REFERENCES

- BRUCE, R. W. Conditions of transfer of training. *J. exp. Psychol.*, 1933, **16**, 343-361.
- BUGELSKI, B. R. Interferences with recall of original responses after learning new responses to old stimuli. *J. exp. Psychol.*, 1942, **30**, 368-379.
- CHENG, N. Y. Retroactive effect and degree of similarity. *J. exp. Psychol.*, 1929, **12**, 444-449.
- DREIS, T. A. Two studies in retroaction: I. Influence of partial identity. II. Susceptibility to retroaction at various grade levels. *J. gen. Psychol.*, 1933, **8**, 157-171.
- GIBSON, E. J. A systematic application of the concepts of generalization and differentiation to verbal learning. *PSYCHOL. REV.*, 1940, **47**, 196-229.
- . Retroactive inhibition as a function of degree of generalization between tasks. *J. exp. Psychol.*, 1941, **28**, 93-115.
- HAMILTON, R. J. Retroactive facilitation as a function of degree of generalization between tasks. *J. exp. Psychol.*, 1943, **32**, 363-376.
- HARDEN, L. M. A quantitative study of the similarity factor in retroactive inhibition. *J. gen. Psychol.*, 1929, **2**, 421-430.
- HOVLAND, C. I. The generalization of conditioned responses: I. The sensory generalization of conditioned responses with varying frequencies of tone. *J. gen. Psychol.*, 1937, **17**, 125-148.
- JOHNSON, L. M. Similarity of meaning as a factor in retroactive inhibition. *J. gen. Psychol.*, 1933, **9**, 377-388.
- KENNELLY, T. W. The role of similarity in retroactive inhibition. *Arch. Psychol.*, N. Y., 1941, **37**, No. 260.
- McGEOCH, J. A. *The psychology of human learning*. New York: Longmans, Green and Co., 1942.
- , & McDONALD, W. T. Meaningful relation and retroactive inhibition. *Amer. J. Psychol.*, 1931, **43**, 579-588.
- , & McGEOCH, G. O. Studies in retroactive inhibition: X. The influence of similarity of meaning between lists of paired associates. *J. exp. Psychol.*, 1937, **21**, 320-329.

15. MCKINNEY, F. Quantitative and qualitative essential elements of transfer. *J. exp. Psychol.*, 1933, **16**, 854-864.
16. MELTON, A. W., & VON LACKUM, W. J. Retroactive and proactive inhibition in retention: evidence for a two-factor theory of retroactive inhibition. *Amer. J. Psychol.*, 1941, **45**, 157-173.
17. OSGOOD, C. E. Meaningful similarity and interference in learning. *J. exp. Psychol.*, 1946, **36**, 277-301.
18. —. An investigation into the causes of retroactive interference. *J. exp. Psychol.*, 1948, **38**, 132-154.
19. ROBINSON, E. S. The 'similarity' factor in retroaction. *Amer. J. Psychol.*, 1927, **39**, 297-312.
20. SIIPOLA, E. M. The relation of transfer to similarity in habit-structure. *J. exp. Psychol.*, 1941, **28**, 233-261.
21. UNDERWOOD, B. J. The effect of successive interpolations on retroactive and proactive inhibition. *Psychol. Monogr.*, 1945, **59**, No. 273.
22. WATSON, B. The similarity factor in transfer and inhibition. *J. educ. Psychol.*, 1938, **29**, 145-157.
23. WICKENS, D. D. The transference of conditioned excitation and conditioned inhibition from one muscle group to the antagonistic muscle group. *J. exp. Psychol.*, 1938, **22**, 101-123.
24. WYLIE, H. H. An experimental study of transfer of response in the white rat. *Behav. Monogr.*, 1919, **3**, No. 16.
25. YUM, K. S. An experimental test of the law of assimilation. *J. exp. Psychol.*, 1931, **14**, 68-82.

[MS. received September 9, 1948]

# THERE IS MORE THAN ONE KIND OF LEARNING<sup>1</sup>

BY EDWARD C. TOLMAN

*University of California*

I wish to suggest that our familiar theoretical disputes about learning may *perhaps* (I emphasize 'perhaps') be resolved, if we can agree that there are really a number of different kinds of learning. For then it may turn out that the theory and laws appropriate to one kind may well be different from those appropriate to other kinds. Each of the theories of learning now current may, in short, still have validity for some one or more varieties of learning, if not for all. But to assume that this will settle our squabbles is, I know, being overly optimistic. Other theorists will certainly not support what I am going to say. Not only will each of them feel that his theory is basic for all kinds of learning, but also each of these others will be sure to object to the general conceptual framework within which my distinctions alone make sense. Thus, whereas I would like to hope that this paper will prove an end to all future papers on learning, I realize that such a hope is mere fantasy or wish-fulfillment on my part or something that my clinical colleagues would undoubtedly dub by some far more unpleasant name.

But, to get down to business; I am going to hold that the connections or relations that get learned can be separated into at least six types. These I shall name as:

1. Cathexes
2. Equivalence Beliefs
3. Field Expectancies
4. Field-Cognition Modes
5. Drive Discriminations
6. Motor Patterns

<sup>1</sup> Address of the Chairman of the Division of General Psychology of the American Psychological Association, Boston, Sept. 7, 1948.

First, let me indicate, briefly, what I mean by each of these six terms and then let me proceed to a more detailed discussion of the conditions and laws for the acquisition, de-acquisition and forgetting of the relations named by each of these six.

1. *Cathexes*. By this term I mean connections or attachments of specific *types* of final positive goal-object, or of final negative 'disturbance-object' to basic drives. (Note that I have coined the term final 'disturbance-object' to cover what have sometimes been called negative goals.) I shall not argue the question as to how many, or what the basic drives may be. I shall assume, however, that you will agree that there are some. For example, none of you will dispute, I hope, the reality of hunger, thirst, sex or fright. By the learning of a cathexis I shall mean, then, the acquisition of a connection between a given variety of goal-object or disturbance-object—*i.e.*, a given type of food, a given type of drink, a given type of sex-object or a given type of fear object—and the corresponding drive of hunger, thirst, sex or fright. That is, the learning of cathexes is the acquisition by the organism of positive dispositions *for* certain *types* of food, drink, sex-object, etc. or of negative dispositions *against* certain *types* of disturbance-object.

2. *Equivalence Beliefs*. This term sounds shocking. However, I am going to use it. By an equivalence belief I mean a connection between a positively cathected type of goal and a type of sub-goal or between a negatively cathected type of disturbance-object and a type of what may be called a sub-disturbance object or *foyer* (*i.e.*, a sort of antecham-

ber which, if the organism gets to it, tends to lead him willy-nilly into the presence of the final disturbance-object itself). During any period in which such an equivalence belief holds, the organism will tend to approach such a type of sub-goal or to avoid such a type of foyer with almost the same readiness with which it will approach a final goal or avoid a final disturbance-object.

3. *Field Expectancies.* These I formerly called 'sign-gestalt-expectations,' which latter term (to quote Gordon Allport (1)), Hilgard and Marquis (10), 'mercifully' shortened to 'expectancies.' This last term is, however, I feel, too disgustingly short; so I am 'mercilessly' rechristening these entities '*field expectancies*.' It is my contention that when an organism is repeatedly presented on successive occasions with an environmental set-up, through which he moves and relative to which he is sensitive, he usually tends to acquire an apprehension not only of each group of immediate stimuli as it impinges upon him but he also tends to acquire a 'set' such that, upon the apprehension of the first group of stimuli in the field, he becomes prepared for the further 'to come' groups of stimuli and also for some of the interconnections or field relationships between such groups of stimuli. It is such sets (or field expectancies) which make it possible for the organism, human or animal, to exhibit appropriate short-cuts and roundabout routes. It is also the acquisition of such sets which make possible the phenomenon of latent learning when (and if) it occurs.

4. *Field-Cognition Modes.* A careful analysis of the processes involved in the appearance of field expectancies indicates, I believe, that the final form and range of any such expectancy is a function not only of repetition, *i.e.*, of 'memory' in the strict sense, but also of 'perception' and of 'inference.' That is, any given field expectancy which appears in

a given experimental set-up is a function of the interacting processes of perception, memory and inference. The modes or manners of functioning of perception, memory and inference are what I am designating as *Field-Cognition Modes*. And I would now assert, further, that in the course of the usual learning experiment there may be acquired not only a specific new field expectancy but also new modes or ways of perceiving, remembering and inferring—new *Field-Cognition Modes* which will, or may, be then utilized by the given organism in still other later environmental set-ups.

5. *Drive Discriminations.* It appears, from some of the latent learning experiments that rats may have to learn to discriminate their thirst from their hunger. In the first Spence and Lippitt (18) experiment the rats were run, when thirsty, with water down one arm of a Y-maze and food down the other. Then, after this preliminary training under thirst, they were shifted to hunger and tested to see if in their free choices they now would immediately choose the food side. They did not. They continued to go to the water side. At first blush this result would be interpreted as a verification of the reinforcement theory. The response of going to the right hand—or water side—had, it will be said, been reinforced by the thirst-reduction which followed the taking of that side during the training trials. It is to be noted, however, that the change from thirst in the preliminary training trials to hunger in the test trials should also have changed (to talk in Hull's language) (12) the  $S_D$  or *Drive-Stimulus*. Further, since (according to Hull) the overt response of turning right or left gets conditioned not only to the maze stimuli but also to this  $S_D$ , this change in  $S_D$  should have caused some breakdown in the learned response. Such a breakdown did not appear. It would seem, therefore, that—speaking in this same lan-

guage—the new hunger  $S_D$  must, under the conditions of this experiment, have for some reason remained undifferentiated from the original thirst  $S_D$ .

However, there are other experiments in which such drive discriminations—to use my language—have proved to be possible and to control the results. Thus, we may recall the Hull (11) and the Leeper (17) experiments in which the animals were run hungry and thirsty on alternate nights. Food was down one alley and water down the other, and both Hull and Leeper found that the animals could learn to turn to the food side when hungry and to the water side when thirsty. Obviously in those experiments, since all the other features were held constant, the two drives—hunger and thirst—*were* discriminated. It appears that this alternation of thirst and hunger and the different locations of the two corresponding rewards throughout the training trials may have been the crucial factor which in these experiments favored such drive discriminations. But there are undoubtedly other ways of inducing drive discriminations, the laws of which a complete psychology of learning must investigate.

6. *Motor Patterns.* It will be noted that this category has to be included by me because I do not hold, as do most behaviorists, that all learning is, as such, the attachment of responses to stimuli. Cathexes, equivalence beliefs, field expectancies, field-cognition modes and drive discriminations are not, as I define them, stimulus-response connections. They are central phenomena, each of which may be expressed by a variety of responses. The actual nature of these final responses is, however, also determined by the character of the motor patterns at the organism's command. My psychology of learning must, therefore, also include a consideration of the laws governing the acquisition of motor patterns purely as such.

So much for a preliminary survey of what I mean by these six terms. Let us turn now to a more detailed discussion of the conditions and laws for the acquisition, de-acquisition and forgetting of each of these six subject-matters for learning.

#### (1) *Cathexes*

In the first place, the distinction between 'positive cathexes' and 'negative cathexes' must be more sharply drawn. By a 'positive cathectis' I mean the attachment of a type of positive goal to a positive drive. That is, when a type of goal has been positively cathected it means that when the given drive is in force the organism will tend to apprehend, to approach, and to perform the consummatory reaction upon any instance of this type of goal which is presented by the immediate environment. By a 'negative cathectis' I mean the attachment of a type of disturbance object to a negative drive. That is to say, when a type of disturbance object has been negatively cathected it means that, if the given negative drive is strong, the organism will tend to apprehend and to avoid or to get away from any instance of this type of disturbance object which is presented in the immediate environment. But let us turn now to the acquisition, the de-acquisition and the forgetting of these two types of cathexes.

##### (a) *Positive cathexes*

It would seem that animals or human beings acquire positive cathexes for new foods, drinks, sex-objects, etc., by trying out the corresponding consummatory responses upon such objects and finding that they work—that, in short, the consummatory reactions to these new objects do reduce the corresponding drives. Hence, here I believe, with Hull (12), in the efficacy of reinforcement or need-reduction. I shall assert, however, that no good experimental evidence has as

yet been adduced for this conclusion. The sort of evidence one wants could be obtained perhaps for hunger, with a special dog-preparation. If, for example, a dog's esophagus were severed and the upper end brought to the outside so that food taken into the mouth, after chewing and swallowing, would drop out into the open and if also a direct fistula were made into the stomach, so that this food (or other food) could then be re-introduced by the experimenter directly into the dog's stomach, we would have the sort of set-up we need. For with this kind of preparation we could discover whether the dog's hunger would become cathected only to those foods which after being chewed and swallowed were re-introduced into the stomach (and hence produced drive-reduction) and whether conversely his hunger would not become cathected to or become de-cathected from foods which in contrast were not re-introduced into the stomach and hence did not produce need-reduction. Furthermore, with such a prepared dog the exact quantitative laws relative to frequency of trials, amounts of reinforcement per trial, etc., could be worked out. Curves could be fitted. Equations for these curves could be mathematically determined and the magnitudes of the constants could be found. In fact, all the precise techniques of quantitative method could be elegantly carried out with such a dog-preparation and bring about closure for all those psychologists who are probably at heart mere physicists or perhaps mathematicians gone wrong. But, prior to such an elegant experiment, all I will say is that I believe that numbers of repetitions and amounts of need-reduction per repetition would, no doubt, turn out to be the two major causal variables and that the curves would undoubtedly be exponential in form.

Next, it must be asked, how is a positive cathectis once acquired, subsequently

de-acquired. How do we come no longer to love specific foods, specific drinks, specific sex-objects? Here I suspect, we have at present even less evidence. If, however, we were to carry out the experiment with our dog preparation, I would suppose that if food were no longer consistently re-introduced into the stomach each time after it had been chewed and swallowed, the cathectis for this type of food would weaken. That is, the failure of reinforcement would, I believe, break the cathectis. And, again, the precise shape and equation for such a curve of de-acquisition could be obtained.

Finally, what about forgetting? Are positive cathectes weakened by the mere passage of time? There seems to be no controlled evidence. But everyday experience suggests that the forgetting of positive cathectes, if it occurs at all, is extremely slow; so that it probably requires years for such cathectes to disappear through mere passage of time and lack of exercise.

#### (b) *Negative cathectes*

The conditions for the acquisition of negative cathectes seem to be well summarized by the ancient adage "a burnt child dreads the fire." In other words, negative reinforcement would seem to be the typical way in which a negative cathectis is acquired. And by negative reinforcement I mean pain or some other type of noxious physiological state.

Conditioning experiments with electric shock as the unconditioned stimulus would seem to provide the model experiments. These experiments, as we now have them, suggest that the curves are steeper for the learning of negative cathectes than for the learning of positive ones. Indeed ordinary experience suggests that a single negative reinforcement may have an overpoweringly strong and persistent effect.

Next, what are the laws for the de-

acquisition of negative cathexes. How do we unlearn our fears? My guess would be that the only way a negative cathexis is broken is by forcing the individual to stay in the presence of the fear-catheted type of object under conditions in which this type of object does not lead to any noxious physiological result.

Finally, as to forgetting; again no controlled evidence. But it seems to me probable that for negative cathexes, as for positive ones, there is practically no forgetting. The same old fears so often seem to endure for a lifetime.

### (2) *Equivalence Beliefs*

By a positive equivalence belief I mean, as I have already indicated, the attachment of a type of sub-goal to a type of final goal such that this sub-goal comes (for the period during which the given equivalence belief holds) to be sought for as if it were the final goal. And by a negative equivalence belief I mean an attachment between a final type of disturbance object and a type of sub-disturbance object or foyer. Now turn to the acquisition, de-acquisition and forgetting of these two types of equivalence beliefs.

#### (a) *Positive equivalence beliefs*

Before discussing these further, a further distinction must be drawn between these beliefs, as such, and the mere apprehensions of objects as appropriate *means* leading on to positive goals. Consider a concrete example—the obtaining of high grades in courses. In so far as high grades are sought merely because they are apprehended as specific means (or paths) leading to the goal, say, of love and respect from teacher or parent, the pursuit of high grades does not in my terms involve an 'equivalence belief.' Operationally speaking, the individual does not stop when he has got the high grade, but he

goes on to use it to obtain the finally wanted love and respect. If, on the other hand, the obtaining of A's in specific courses seems to bring at once, and by itself, some reduction of the underlying drive or drives, then there is involved some strength of an equivalence belief in my sense of that term. That is to say, the individual then accepts what was originally a mere *means* as *equivalent* to a goal. He experiences, when reaching this means, some degree of drive-reduction. The precise observations necessary for determining such phenomena are, however, difficult.

Let us imagine first a case with rats. Suppose, it were found that when rats reached a well-practiced type of goal-box (even though now there is no food in it), their stomach contractions or some more basic physiological measure of hunger subsided at least for a time, then I would say that we had evidence for some degree of equivalence belief in these rats to the effect that the given type of goal-box was equivalent to food.

Equivalence beliefs will, however, I believe, get most frequently established, not in connection with such a viscerogenic drive as hunger, but in connection with social drives. I cannot here attempt to argue for the validity and reality of such drives. I hope merely that for the purposes of this discussion you will grant me them. If you will, let us consider again the example of the student working for high grades. In so far as it can be demonstrated that with the reception of the high grades there is some temporary reduction in this student's need for love and approbation, even without his going on to tell others about his grade, then we would have evidence for an equivalence belief. The A's would then be accepted by him as equivalent to the love or approbation to which they were originally a mere means. The difficulty is, of course, that we have no good techniques for measuring the

varying moment-by-moment strengths of any such drives as the need for love or approbation. Some day, however (perhaps by an improvement of projective techniques), we may acquire such a method. And then, as with hunger, we can see if the drive does actually subside in the mere presence of the high grade.

But what would be the laws for the acquisition, de-acquisition and forgetting of such equivalence beliefs? Tentatively, I would propose primacy, frequency and intensity of need-reductions plus early traumatic experiences as the important causal factors. That is, I would hold that the earliness (or primacy) as well as the frequency of the occasions in the life of the individual in which getting high grades, led to love would be important. I would also suggest that early traumatic accompaniments of this sequence between grades and social approbation would also help to 'fixate' the getting of high grades as an end in themselves. But as to the laws or shapes of the curves of acquisition I am completely in the dark. This is a virgin field in which the clinically minded experimentalist and the experimentally minded clinician might well cultivate conjointly.

And, as to the laws for the de-acquisition of equivalence beliefs, I am equally in the dark. And this is sad, because I would assert that a large part of clinical practice consists in the attempt to break erroneous equivalence beliefs. As long as an equivalence belief is not misleading in the sense that the sub-goal is usually actually followed by the true goal, the belief probably serves some physiological economy. But when an equivalence belief persists, even though now the sub-goal practically never leads to the true goal, such an equivalence belief would seem bad. The subject experiences some drive-reduction (Query: so-called 'secondary gain') but most of his actual drive remains unsatisfied. He

accepts the shadow for the substance. He continues to be a greasy grind, although now as a young adult few praise him for it and many even contemn him. And so we find him at last rushing up to the University Psychiatric Counseling Center, or whatever locally it may be called, to find out why he is so anxious and/or so depressed.

Our question becomes how actually do the clinicians, the psychiatrists, the counselors (or whatever they may be called) break such erroneous beliefs. What are the conditions and laws governing their de-acquisition? Nobody seems to know. The therapists do have considerable success. But any clear, agreed-upon statement of their procedure seems to be lacking. Some talk about the 'transference relation.' Others say that the patient is really a 'feeling sensation' type and that he has been trying to operate as if he were a 'thinking intuition' type. Others 'reflect back' to the patient what he himself says. And some merely give him a good sound lecture and tell him to go about his business. But what is common in all these procedures and why they all in some degree succeed in breaking erroneous equivalence beliefs seems still quite beyond us (or at any rate beyond me).

Finally, as to the forgetting or non-forgetting of positive equivalence beliefs through a mere passage of time, we seem to have little evidence. But general experience suggests that such beliefs are not merely forgotten but have to be unlearned.

#### (b) *Negative equivalence beliefs*

It must be pointed out that negative equivalence beliefs cannot really be differentiated from negative cathexes. For example, the rat is conditioned against the box which led to shock. He also comes to be conditioned against the type of paths which lead to such boxes. Or again a human individual is shown to

have developed a tendency to avoid self-assertive behaviors. And it turns out that such self-assertive behaviors were followed in childhood by parental disapproval, and parental disapproval was followed by a final physiological disturbance. The individual in question has established an equivalence belief that such self-assertive behavior was equivalent to loss of love and resultant physiological disturbance. And, though when this man grows up, this belief may no longer be correct (his parents may be dead or may in fact like in the man what they punished in the child), he nevertheless continues to avoid self-assertive behaviors in situations in which, instead of such behaviors being punished, they would actually be rewarded.

So once again the problem of the laws and the conditions for the acquisition, the de-acquisition and the forgetting of equivalence beliefs becomes clinically important. But again we have no clear evidence. Are negative equivalence beliefs more rapidly established than positive ones? What are the therapeutic procedures most favorable to their de-acquisition? Are they ever forgotten as a result of mere non-exercise? Let me suggest again that the experimentally minded therapist and the clinically oriented experimentalist get together and find out.

### (3) *Field Expectancies*

Here my notions as to the conditions and laws of learning depart perhaps most radically from those ordinarily held. By field expectancies I mean (to recall what I said above) those sets which get built up in an organism relative to a specific environmental field. These are the sets which, after learning, make it possible for the organism not only to choose correctly the particular paths in the field on which he has been practiced but also, in some degree, to perform correctly on short-cuts and

roundabouts not previously practiced. It is, of course, the facts of latent learning plus the facts of taking short-cuts and roundabouts, when forced or permitted, which have driven me and others to the notion that when a rat, or a human being, is practiced in a particular set of activities in a particular environment, an essential part of what he acquires is an expectancy, a sign-Gestalt, a cognitive structure, a cognitive map (to use some of the terms which have been suggested) relative to that environment.

Further (and here I confess I have up to now been somewhat unclear) I used to be so impressed by the latent learning experiments of the type invented by Blodgett (3), in which no reward was introduced during the learning period, that I was apt to formulate the conditions involved in such field-expectancy learning primarily in terms of frequency alone and as if motivation played no role. However, if I did this, I was in error. It is obvious that completely unmotivated animals will not learn. They will go to sleep or otherwise divorce themselves from the task. So it must be emphasized that in the Blodgett experiments, even though the animals were not rewarded, they *were* motivated. Also, from the first Spence and Lippitt experiment (18) and from some of the follow-up experiments by Kendler (13) and Walker (19), it appears that mere exercise, mere exploration under one drive, may not be enough to cause the animals to perceive and to learn the position of reward-objects appropriate to some other drive. Thirsty animals apparently do not notice food, even though the experiment be rigged as it was by Kendler and Mencher (14) to seem to force them to notice that the cups which did not contain water did contain food. Summing it up then, it appears that motivation conditions *are* very important for the building up of field-expectancies. I

would like in this connection to report briefly an experiment recently done by Gleitman at California. He used a T-maze and trained hungry rats to get equal amounts of food at each end; the two end boxes being quite dissimilar in character. Then these two end-boxes were placed in another room and the rats were introduced into each of them. In one they received a shock and in the other no shock. They were then immediately replaced in the original maze, and 22 out of 25 animals immediately avoided at the choice point the pathway which led to the end-box in which they had just been shocked. This showed that during the previous training they *had* learned which path at the choice point led to which end-box in spite of the fact that they had been equally reinforced in both end-boxes. In other words, their hunger, as well as their exercise had probably led them under these conditions to build up spatial sign-Gestalten which could now be appropriately used for a different response, namely for that of now avoiding the end-box in which punishment had just been received. That is, rats can learn under hunger which path leads to which end-box and they can learn that a given end-box now means punishment and *not* food—even though they apparently cannot perceive water and learn its location when under strong hunger, nor perceive food and learn its location when under strong thirst.

To sum up, I would conclude that motivation conditions must be assumed to play a role in the building up of field expectancies. But this does not mean that I hold that such learning consists in the stamping in of S-R habits by reinforcement. The presence of reinforcement in a particular locus makes that locus a goal which determines what performance will take place but it does not stamp in S-R connections though it

probably does give a special vividness to that locus in the total field expectancy.

The main question is, however, what are the laws determining the acquisition, de-acquisition and forgetting of such field expectancies.

As to acquisition.—First we have to know, for the given species, the facts of their perceptual sensitivity. Obviously the field expectancies, which get built up, can include only such aspects of the environment as the given organism is capable of perceiving.

Secondly, we have to know the facts concerning the ability of the given organism (under the given conditions of motivation) to connect and associate the different parts of the field so that when he is in one part of the field he will *remember* what was present in other parts.

Thirdly, we have to know the facts concerning what, for want of a better name, we may call the animal's '*inference abilities*.' These would state the capacity of the given individual, or species, to extend its expectancies *re* given environmental fields beyond the parts upon which this individual has been specifically exercised. It is these capacities which will underlie the animal's ability to short-cut and to take *Umwege*. Such inference facts, when we have them, will obviously be found to include something about an ability on the animal's part to set up a system, or systems, of orienting coordinates as a result of the presence, or absence, of such and such strategically located cues. All this, of course, sounds complicated. But personally, it seems to me that we are a very long way from any precise laws for such field-expectancy learning, and where we seem to have such laws it is because either overtly (or covertly) we have held constant most of the important circumambient variables. We can work out equations and constants for the development of specific behaviors in specific apparatuses under specific

motivations. But how the form of the specific apparatus plus the nature and magnitudes of the specific motivation enter into and determine these equations, I believe we do not know.

Granted then, that we are still very near the beginning of our knowledge of laws for the acquisition of field expectancies, what can we say about the laws of their de-acquisition and forgetting? Here, I have nothing but simple hunches to offer. These hunches would be, first, that the de-acquisition of field expectancies only takes place when the actual environment is so changed that the previous expectancy is no longer suitable. The de-acquisition of one field expectancy results from the learning of another conflicting expectancy. But to what extent the laws and equations for such *new learnings* will be different from those for the original learnings, I hesitate to say. I should expect the equations to have the same form but that new constants would be required.

Finally, as to forgetting, here in the case of field expectancies, as contrasted with that of cathexes and of equivalence beliefs, I believe that true forgetting (*i.e.*, weakening as a result of the mere passage of time) *does* take place. We don't forget our cathexes and we don't forget our equivalence beliefs, but we do forget particular environmental lay-outs which we have not experienced for long periods of time—though this forgetting obviously obeys the sorts of laws which the Gestalt psychologists have uncovered and not the old simple associationistic ones. The remembered environmental lay-out becomes changed, *i.e.*, simplified or sharpened, as well as weakened by the mere passage of time. Some features become enhanced, others minimized or even dropped out and some wholly new features may be added. The work of the Gestalt psychologists (see for example, Koffka, 15 and of Bartlett 2) all bear eloquent testimony to such non-

associationistic features in the forgetting of field-expectancies.

Turn now to the next category.

#### (4) *Field-Cognition Modes*

This category is the one about which I am least confident. Perhaps a better name would be field lore—that is, perceptual, memorial and inferential lores. Much of these lores—particularly perceptual lore—seems to be given innately. It is the lore that such and such stimulus configurations are to be taken as cues for such and such perceptions. And although a solid basis of such lore is nativistically given, it can, as we know, be added to and modified by experience. Consider, for example, the experiment of Fieandt (one of Brunswik's students, 6) in which the subjects came, as a result of training, to use a very slight mark on one of each pair of cards, as a cue which led them to perceive this card as having an objectively light shade but as in shadow. The slight mark, often not perceived consciously as such, came nonetheless to be used as a perceptual cue that the card was in shadow and therefore really light. And in so far as this tendency would transfer to new situations, it would be an example of an acquired perceptual lore.

Memorial lore is relatively simple. The one innately given principle seems to be that, if a certain sequence of events has occurred on one occasion, this same sequence of events is likely to occur on subsequent occasions. This principle seems to be innately strong. The complementary principle of mere probabilities of occurrence is one which has to be learned. Brunswik (4) more than any other psychologist has investigated this question of the learning to expect mere probabilities. But even he, I think, has not carried the implications far enough. Thus, for example, it would seem to me that the ability to 'tolerate ambiguity,' a concept developed by Else

Frenkel-Brunswik (7), probably closely ties in with this learning to expect mere probabilities. But the conditions of early childhood training, or whatever they may be, which develop this new memorial principle that allows the subject to be able to remember not 100 per cent sequences but merely probable sequences, that is, to tolerate ambiguities, have yet to be subjected to more study.

Inferential lore, in its simplest form, would like perceptual lore and memorial lore be to a considerable extent innate. It would consist of the simple rules of space, time, force, and quantity, the bases for which are certainly innate. Such lore receives, however, tremendous additions through specific verbal training—especially in us human beings. We men learn verbally all sorts of rules about time, space, force and quantity. And these rules we then carry around with us from one specific situation to another—so that they then underlie and govern our specific apprehensions, *i.e.*, our field expectancies, for each new environmental field.

In a word, I am trying to summarize under this fourth category all those principles as to the structure of environmental fields which are relevant to all environmental fields, and which (whether innate or learned) are carried around by the individual and applied to each new field with which he is presented.

As to the conditions and laws for the acquisition, de-acquisition and forgetting of such perceptual, memorial and inferential modes, as distinct from the acquisition of the concrete apprehension of the particular fields themselves, I believe we have as yet practically no information.

Turn now to the fifth category.

#### (5) *Drive Discriminations*

Here, also, I have but little more to say. In my preliminary remarks con-

cerning this category I referred to the latent learning experiments and to the Hull and the Leeper experiments which suggested that rats sometimes may have to learn to distinguish between their different drives. And I believe that there are similar learnings required of human beings. We, too, I believe, often have to learn to discriminate our true needs. In fact, I would suggest that sometimes the task of psychotherapy is not merely, as I argued above, that of breaking incorrect, yet traumatically held to, equivalence beliefs. It may also be, on occasion, the helping of the patient to learn to discriminate his real drives or needs.

But, again, we have practically no experimental data either for rats or for men as to how we learn, unlearn, or forget (if we do) these drive discriminations.

#### (6) *Motor Patterns*

Guthrie (8) has emphasized the learning of 'movements.' Where 'movements' are contrasted with 'acts.' Acts he admits to be goal directed (although in the last analysis it would appear that for him they also must dissolve into complexes of non-goal directed movements). But, in any event, in calling attention to movements Guthrie is calling attention, I think, to what I would mean by motor patterns. And in default of other experimental theories about the learning of motor patterns I am willing to take a chance and to agree with Guthrie that the conditions under which a motor pattern gets acquired may well be those in which the given movement gets the animal away from the stimuli which were present when the movement was initiated. Any response (*i.e.*, any movement) which goes off will, according to Guthrie, get conditioned on a single trial to whatever stimuli were then present. Therefore a movement which removes the individual from out the range of those stimuli tends to be the one which

remains conditioned to them because no other movements have a chance to occur and to displace it. A motor pattern thus gets learned without reinforcement. I would like to point out, however, that such a learning of motor patterns is of necessity always imbedded in a larger goal-directed activity—a point which is not emphasized by Guthrie. His and Horton's cats (9) did learn stereotyped motor patterns for getting out of their hit-the-barber-pole type of problem-box; but they learned them only because they, the cats, were involved in the larger goal-directed activity of getting to the food in front. And, similarly, I believe that rats learn stereotyped motor patterns for running specific mazes only when these specific patterns actually get them to food. When such specific movements do not succeed, trial and error supervene and new movements get a chance to become conditioned, but again only if these new ones prove in the larger setting to get the animal to his goal.

Finally, however, once a movement sequence gets learned in one situation, it is ready, I believe, to be tried out in other situations. We do build up, I believe, many motor patterns (the old name was sensory-motor skills) which we carry around with us as equipment for behaving in new situations. And, whereas, I do not think we as yet know much about the laws for the learning, unlearning and forgetting of such motor patterns, I am willing to accept, for the present, Guthrie's notions concerning their learning and unlearning. Finally, as to their forgetting I can merely point to the everyday fact that one's skills do seem to get rusty with lack of exercise and the passage of time.

*Now to conclude:* Let me briefly summarize. There are, I believe, at least six kinds of learning—or rather the learning of at least six kinds of relationship. I have called these six relationships: cathexes, equivalence beliefs, field

expectancies, field-cognition modes, drive discriminations and motor patterns. And, although, as usual, I have been merely programmatic and have not attempted to set up, at this date, any precise systems of postulates and deduced theorems, I *have* made some specific suggestions as to some of the conditions and laws for the acquisition, de-acquisition and forgetting of these relationships. I feel that once we have thought of really good defining experiments for each of these types of learning we can then hypothesize equations, fit empirical curves and dream up constants to our hearts' content. At least I *think I could*.

Summarizing more specifically for each of the six kinds of learning the following further suggestions were also made. (1) I suggested that the 'reinforcement' doctrine is probably valid for the acquisition of cathexes. (2) I suggested that this 'reinforcement' principle plus traumatic experience is probably also valid for the acquiring of equivalence beliefs. And I asserted that erroneous equivalence beliefs are a large part of what the therapist has to contend with. (3) I held that reinforcement *per se* is not valid for the acquisition of field expectancies and I also emphasized that Gestalt principles of learning and forgetting, rather than associationistic principles, are of prime importance in the acquiring and the forgetting of such field expectancies. (4) For the acquisition, de-acquisition and forgetting of the field-cognition modes of perception, memory, and inference I had no laws to suggest. The development of such laws would depend upon the carrying out of many more carefully designed transfer experiments than we now have. (5) For the learning of drive discriminations I also had no laws but pointed to types of experiment which seem to have favored the development of drive discriminations. (6) Finally, as to the laws for the acquisition of motor patterns, *per se*, I suggested that Guth-

rie's principle of simple conditioning may perhaps be correct.

*One last word.* Why do I want thus to complicate things; why do I not want one simple set of laws for all learning? I do not know. But I suppose it must be due to some funny erroneous equivalence belief on my part to the effect that being sweeping and comprehensive, though vague, is equivalent to more love from others than being narrow and precise. No doubt, any good clinician would be able to trace this back to some sort of nasty traumatic experience in my early childhood. Let, then, the clinician unravel this sort of causal relationship in me or in others and I will attempt to show him its analogue in rats, or at least in chimpanzees or perhaps dogs. For, if more of the theoretical and learning psychologists, on the one hand, and of the clinicians, on the other, don't get together, and soon, there is really going to develop that nasty fission in psychology (5, 16) that we have all been warned of. And, if that fission happens, then our science is really going to suffer a long and very unfortunate period of schizophrenic "institutionalization"—whether inside of or outside of our universities.

#### REFERENCES

1. ALLPORT, G. W. Scientific models and human morals. *PSYCHOL. REV.*, 1946, **54**, 182-192.
2. BARTLETT, F. C. *Remembering—a study in experimental and social psychology*. Cambridge: University Press, 1932.
3. BLODGETT, H. C. The effect of the introduction of reward upon the maze performance of rats. *Univ. Calif. Publ. Psychol.*, 1929, **4**, 113-134.
4. BRUNSWIK, E. Probability as a determiner of rat behavior. *J. exp. Psychol.*, 1939, **25**, 175-197.
5. CRANNELL, C. W. Are rat psychologists responsible for fission? *Amer. Psychologist*, 1947, **2**, 22-23.
6. FIEANDT, KAIV. Dressurversuche an der Farbenwahrnehmung. *Arch. f. d. ges. Psychol.*, 1936, **96**, 467-495.
7. FRENKEL-BRUNSWIK, E. Tolerance toward ambiguity as a personality variable. *Amer. Psychologist*, 1948, **3**, p. 268.
8. GUTHRIE, E. R. Association and the law of effect. *PSYCHOL. REV.*, 1940, **47**, 127-148.
9. —, & HORTON, G. P. *Cats in a puzzle box*. New York: Rinehart & Company, Inc., 1946.
10. HILGARD, E. R. & MARQUIS, D. G. *Conditioning and learning*. New York: D. Appleton-Century Company, 1940.
11. HULL, C. L. Differential habituation to internal stimuli in the albino rat. *J. comp. Psychol.*, 1933, **16**, 225-273.
12. —. *Principles of behavior*. New York: D. Appleton-Century Company, 1943.
13. KENDLER, H. H. An investigation of latent learning in a T-maze. *J. comp. physiol. Psychol.*, 1947, **40**, 265-270.
14. —, & MENCHER, H. C. The ability of rats to learn the location of food when motivated by thirst—an experimental reply to Leeper. *J. exp. Psychol.*, 1948, **38**, 82-88.
15. KOFFKA, K. *Principles of Gestalt psychology*. New York: Harcourt, Brace and Company, 1935.
16. KRECH, D. A note on fission. *Amer. Psychologist*, 1946, **1**, 402-404.
17. LEEPER, R. The role of motivation in learning: a study of the phenomenon of differential motivational control of the utilization of habits. *J. genet. Psychol.*, 1935, **46**, 3-40.
18. SPENCE, K. W., & LIPPITT, R. An experimental test of the sign-gestalt theory of trial and error learning. *J. exp. Psychol.*, 1946, **36**, 491-502.
19. WALKER, E. L. Drive specificity and learning. *J. exp. Psychol.*, 1948, **38**, 39-49.

[MS. received October 7, 1948]

## THE GESTALT THEORY OF EXPRESSION

BY RUDOLF ARNHEIM

*Sarah Lawrence College*

What is the exact location and range of the territory covered by the term 'expression'? Thus far, no generally accepted definition exists. In order to make clear what is meant by expression in the present paper, it is therefore necessary to indicate (1) the kind of perceptual stimulus which involves the phenomenon in question, and (2) the kind of mental process to which its existence is due. This delimitation of our subject will show that the range of perceptual objects which carry expression according to gestalt theory is unusually large and that expression is defined as the product of perceptual properties which various other schools of thought consider non-existent or unimportant.

(1) In present-day usage, the term 'expression' refers primarily to behavioral manifestations of the human personality. The appearance and activities of the human body may be said to be expressive. The shape and proportions of the face or the hands, the tensions and the rhythm of muscular action, gait, gestures, and other movements serve as objects of observation. In addition, expression is now commonly understood to reach beyond the observed person's body. The 'projective techniques' exploit characteristic effects upon, and reactions to, the environment. The way a person dresses, keeps his room, handles the language, the pen, the brush; the colors, flowers, occupations he prefers; the meaning he attributes to pictures, tunes, or inkblots; the story he imposes on puppets; his interpretation of a dramatic part—these and innumerable other manifestations can be called 'expressive' in that they permit conclusions about the personality or the temporary

state of mind of the individual. Gestalt psychologists extend the range of expressive phenomena beyond this limit. For reasons which will be discussed, they consider it indispensable to speak also of the expression conveyed by inanimate objects, such as mountains, clouds, sirens, machines.

(2) Once the carrier of expression is determined, the kind of mental process must be indicated which is charged with producing the phenomenon. It is the contention of Gestalt psychology that the various experiences commonly classified under 'perception of expression' are caused by a number of psychological processes, which ought to be distinguished from each other for the purpose of theoretical analysis. Some of these experiences are partly or wholly based upon empirically acquired knowledge. The mere inspection of many half-smoked cigarettes in an ashtray would suggest no connection with nervous tension to a visitor from a planet inhabited by non-smokers. The letters EVVIVA GUERRA and EVVIVA DON PIO scribbled all over the walls of an Italian village will reveal the mentality of the natives only to someone who happens to know that these words pay homage to a champion cyclist and the village priest. For the purpose of the present paper, the use of past experience for the interpretation of perceptual observations will be excluded from the field of expression and referred to the psychology of learning. We shall be concerned only with instances in which, according to Gestalt psychology, sensory data contain a core of expression that is perceptually self evident. The way a person keeps his lips tightly closed or raises his voice or

strokes a child's head or walks hesitatingly is said to contain factors whose meaning can be understood directly through mere inspection. Instances of such direct expression are not limited to the appearance and behavior of the subject's own body. They are also found in such 'projective' material as the stirring red of a woman's favorite dress or the 'emotional' character of the music she prefers. In addition, inanimate objects are said to convey direct expression. The aggressive stroke of lightning or the soothing rhythm of rain impress the observer by perceptual qualities which according to Gestalt psychology must be distinguished theoretically from the effect of what he knows about the nature of these happenings. It is assumed, however, that practically every concrete experience combines factors of both kinds.

*Procedures and findings.* What is expression, and what enables the observer to experience it? By means of which perceptual factors and in what way do stimulus configurations evoke such experiences in the onlooker? During the last twenty-five years or so, numerous experimental investigations have been devoted to the phenomena of expression, but hardly any of them have tried to answer our questions. Limited as they were to the connection between how a person behaves and what happens in him psychologically, they centered upon the certainly important problem: To what extent are observers, untrained or trained, gifted or average, capable of getting valid information about a person's temporary state of mind or his more permanent psychical constitution from an inspection of his face, voice, gait, handwriting, etc.?

This is true for the various matching-experiments, which are conveniently summarized by Woodworth (24, pp. 242-256) and by Allport and Vernon (1, pp. 3-20). Similarly, in the field of

the projective techniques psychologists have looked for correlations between personality traits and reactions to environmental stimuli. Almost invariably, these stimuli contain factors of the kind which concern the present paper. However, thus far, little explicit discussion has been devoted to the question why and how the given percepts provoke the observed reactions. There is evidence that the whole structure of a face rather than the sum of its parts determines expression (2). But which structural features make for what expression and why? In the Rorschach test, the typical reactions to color are probably based on expression. But why are emotional attitudes related to color rather than shape? Ernest G. Schachtel has done pioneer work in this field, pointing out, for instance, that responses to colors and to affect-experiences are both characterized by passive receptivity (19). On the whole, however, questions of this kind have been answered thus far by summary and scantily supported theoretical assertions.

A few remarks are in order on the investigations which have tested the accomplishments of observers. A glance at the results reveals a curious contrast. One group of experimenters reports essentially negative findings. Another, consisting mainly of Gestalt psychologists, asserts that observers judge portraits, handwritings, and similar material with a measure of success that clearly surpasses chance. Pessimistic generalizations have been drawn from the studies of the first type. The subject of expression is sometimes treated with the buoyant unkindness that distinguished the early behavioristic statements on introspection. This attitude has not encouraged research.

The main reason for the conflicting results can be found in differences of approach. The investigators of the first type asked: How validly can the bodily

expression of the average person or of a random member of a particular group of people be interpreted? In other words, they focused on the important practical question of the extent to which expression can be relied upon in everyday life. On the other hand, the Gestalt psychologists preferred the common scientific procedure of purifying as carefully as possible the phenomenon under investigation. They searched for the most favorable condition of observation. A major part of their efforts was spent in selecting and preparing sets of specimens which promised to demonstrate expression clearly and strongly (2, p. 8).

Some of the factors which may account for the often disappointing results obtained in experiments with random material are the following. (a) Everyday observation suggests that the structural patterns of character, temperament, mood, are not equally clear-cut in all people. While some individuals are pronouncedly depressed or lighthearted, strong or weak, harmonious or disharmonious, warm or cold, others strike us as indefinite, lukewarm, fluid. Whatever the exact nature of such indefiniteness, one would expect the corresponding faces, gestures, handwritings to be equally vague in form and therefore in expression. When one examines material of this kind, one notices in some cases that the decisive structural features are not sharply defined. In other cases, factors which are clear-cut in themselves add up to something that shows neither harmony nor conflict but a lack of unity or relatedness, which renders the whole meaningless, inexpressive. Many telling examples can be found among the composite faces made up by the summation of unrelated parts for experimental purposes. If observers can cope with such material at all, they do so presumably by guessing what these artifacts are meant to mean rather than by having the experience of live

expression. (b) The presence of a portrait photographer's camera tends to paralyze a person's expression, and he becomes self-conscious, inhibited, and often strikes an unnatural pose. (c) Candid shots are momentary phases isolated from a temporal process and a spatial context. Sometimes they are highly expressive and representative of the whole from which they are taken. Frequently they are not. Furthermore the angle from which a shot is made, the effect of lighting on shape, the rendering of brightness and color values, as well as modifications through retouching, are factors which make it impossible to accept a random photograph as a valid likeness. (d) If for purposes of matching experiments a number of samples is combined at random, accidental similarities of expression may occur, which will make distinction difficult, even though every specimen may be clear-cut in itself. Further reasons for the lack of consistent results are discussed by Wolff (23, p. 7).<sup>1</sup>

The conclusion seems to be that the recognition of expression has been proven to be reliable and valid under optimal conditions. For the average face, voice, gesture, handwriting, etc., the results are likely to be less positive. However, in order to establish this fact trustworthy, the additional obstacles

<sup>1</sup> Since there is no reason to expect that every photograph will reproduce essential features of expression, it would be interesting to know by which criterion the photographs for the Szondy-test (18) have been selected. If an integral feature of the test consists in establishing the reactions of people to the personalities of homosexuals, sadistic murderers, etc., two questions arise. (1) Is there a complete correlation between these pathological manifestations and certain clear-cut personality structures? (2) Are the latter suitably expressed in the photographs? These problems are avoided if the test is meant simply to investigate people's responses to a given set of portraits, whatever their origin.

created by unsuitable experimental conditions will have to be reduced.

*Associationist theories.* What enables observers to judge expression? The traditional theory, handed down to our generation without much questioning, is based on associationism. In his essay on vision Berkeley (4, § 65) discusses the way in which one sees shame or anger in the looks of a man.

"Those passions are themselves invisible: they are nevertheless let in by the eye along with colours and alterations of countenance, which are the immediate object of vision, and which signify them for no other reason than barely because they have been observed to accompany them: without which experience, we should no more have taken blushing for a sign of shame, than of gladness."

Darwin, in his book on the expression of emotions, devoted a few pages to the same problem (7, pp. 356-359). He considered the recognition of expression to be either instinctive or learned. "Children, no doubt, would soon learn the movements of expression in their elders in the same manner as animals learn those of man," namely, "through their associating harsh or kind treatment with our actions."

"Moreover, when a child cries or laughs, he knows in a general manner what he is doing and what he feels; so that a very small exertion of reason would tell him what crying or laughing meant in others. But the question is, do our children acquire their knowledge of expression solely by experience through the power of association and reason? As most of the movements of expression must have been gradually acquired, afterwards becoming instinctive, there seems to be some degree of *a priori* probability that their recognition would likewise have become instinctive."

In Darwin's view, the relationship between expressive bodily behavior and the corresponding psychical attitude was merely causal. Expressive gestures were

either remnants of originally serviceable habits or due to 'direct action of the nervous system.' He saw no inner kinship between a particular pattern of muscular behavior and the correlated state of mind.

A variation of the associationist theory contends that judgments of expression are based on stereotypes. In this view, interpretation does not rely on what belongs together according to our spontaneous insight or repeated observation but on conventions, which we have adopted ready-made from our social group. We have been told that aquiline noses indicate courage and that protruding lips betray sensuality. The promoters of the theory generally imply that such judgments are wrong, as though information not based on first-hand experience could never be trusted. Actually, the danger does not lie in the social origin of the information. What counts is that people have a tendency to acquire simply structured concepts on the basis of insufficient evidence, which may have been gathered first-hand or second-hand, and to preserve these concepts unchanged in the face of contrary facts. While this may make for many onesided or entirely wrong evaluations of individuals and groups of people, the existence of stereotypes does not explain the origin of physiognomic judgments. If these judgments stem from tradition, what is the tradition's source? Are they right or wrong? Even though often misapplied, traditional interpretations of physique and behavior may still be based on sound observation. In fact, perhaps they are so hardy because they are so true.

*Empathy.* The theory of empathy holds an intermediate position between the traditional and a more modern approach. This theory is often formulated as a mere extension of the association theory, designed to take care of the expression of inanimate objects. When I

look at the columns of a temple, I know from past experience the kind of mechanical pressure and counterpressure that occurs in the column. Equally from past experience I know how I should feel myself if I were in the place of the column and if those physical forces acted upon and within my own body. I project my feelings into the column and by such animation endow it with expression. Lipps, who developed the theory, stated that empathy is based on association (16, p. 434). It is true, he also says, that the kind of association in question connects "two things belonging together, or being combined by necessity, the one being immediately given in and with the other." But he seems to have conceived of this inner necessity as a merely causal connection, because immediately after the statement just quoted he denies explicitly that the relationship between the bodily expression of anger and the angry person's psychical experience could be described as an "association of similarity, identity, correspondence" (p. 435). Like Darwin, Lipps saw no intrinsic kinship between perceptual appearance and the physical and psychological forces 'behind' it. However, he did see a structural similarity between physical and psychological forces in other respects. After discussing the mechanical forces whose existence in an inanimate object is inferred by the observer through past experience, Lipps writes the following remarkable passage:

"And to (the knowledge of these mechanical forces) is furthermore attached the representation of possible internal ways of behavior of my own, which do not lead to the same result but are of the same character. In other words, there is attached the representation of possible kinds of my own activity, which in an analogous fashion, involves forces, impulses, or tendencies,

freely at work or inhibited, a yielding to external effect, overcoming of resistance, the arising and resolving of tensions among impulses, etc. Those forces and effects of forces appear in the light of my own ways of behavior, my own kinds of activity, impulses, and tendencies and their ways of realization" (16, p. 439).

Thus Lipps anticipated the Gestalt principle of isomorphism for the relationship between the physical forces in the observed object and the psychical dynamics in the observer; and in a subsequent section of the same paper he applies the 'association of similarity of character' even to the relationship between the perceived rhythm of musical tones and the rhythm of other psychical processes that occur in the listener. Which means that in the case of at least one structural characteristic, namely rhythm, Lipps realized a possible inner similarity of perceptual patterns and the expressive meaning they convey to the observer.

*The Gestalt approach.* The Gestalt theory of expression admits that correspondences between physical and psychical behavior can be discovered on the basis of mere statistical correlation but maintains that repeated association is neither the only nor the common means of arriving at an understanding of expression. Gestalt psychologists hold that expressive behavior reveals its meaning directly in perception. The approach is based on the principle of isomorphism; according to which processes which take place in different media may be nevertheless similar in their structural organization. Applied to body and mind, this means that if the forces which determine bodily behavior are structurally similar to those which characterize the corresponding mental states, it may become understandable why psychical meaning can be read off directly from a person's appearance and conduct.

It is not the aim of this paper to prove

TABLE 1  
ISOMORPHIC LEVELS

A. Observed Person		
I. State of mind	psychological	
II. Neural correlate of I	electro-chemical	
III. Muscular forces	mechanical	
IV. Kinesthetic correlate of III	psychological	
V. Shape and movement of body	geometrical	
B. Observer		
VI. Retinal projection of V	geometrical	
VII. Cortical projection of VI	electro-chemical	
VIII. Perceptual correlate of VII	psychological	

the validity of the Gestalt hypothesis.<sup>2</sup> We shall limit ourselves to pointing out some of its implications. Only brief presentations of the theory are available so far. However, Köhler's (12, pp. 216-247) and Koffka's (10, pp. 654-661) remarks about the subject are explicit enough to indicate that isomorphism on only two levels, namely the psychical processes which occur in the observed person and the corresponding behavioral activity, would be insufficient to explain direct understanding of expression through perception. In the following an attempt will be made to list a number of psychological and physical levels, in the observed person and in the observer, at which isomorphic structures must exist in order to make the Gestalt explanation possible.

Let us suppose that a person A performs a 'gentle' gesture, which is experienced as such by an observer B. On the basis of psycho-physical parallelism in

its Gestalt version it would be assumed that the tenderness of A's feeling (Table 1, level I) corresponds to a hypothetical process in A's nervous system (level II), and that the two processes, the psychical and the physiological, are isomorphic, that is to say, similar in structure.

The neural process will direct the muscular forces which produce the gesture of A's arm and hand (level III). Again it must be assumed that the particular dynamic pattern of mechanical action and inhibition in A's muscles corresponds structurally to the configuration of physiological and psychical forces at the levels II and I. The muscular action will be accompanied with a kinesthetic experience (level IV), which again must be isomorphic with the other levels. The kinesthetic experience need not always take place and is not strictly indispensable. However, the structural kinship of the experienced gentleness of his gesture and the equally experienced gentleness of his mood will make A feel that his gesture is a fitting manifestation of his state of mind.

Finally, the muscular forces of level III will cause A's arm and hand to move in a, say, parabolic curve (level V); and again the geometric formation of this curve would have to be isomorphic with the structure of the processes at the previous levels. An elementary geometrical example may illustrate the meaning of this statement. Geometrically, a circle is the result of just one structural condi-

<sup>2</sup> For that purpose, observations of infants are relevant. Even in his day, Darwin was puzzled by the fact that young children seemed directly to understand a smile or grief "at much too early an age to have learnt anything by experience" (7, p. 358). According to Bühler (6, p. 377), "the baby of three or four months reacts positively to the angry as well as to the kind voice and look; the five-to-seven-months-old baby reflects the assumed expression and also begins to cry at the scolding voice and threatening gesture" on the basis of "direct sensory influence." Further evidence will have to come from detailed demonstrations of structural similarities. (Cf. p. 169.)

tion. It is the locus of all points that are equally distant from one center. A parabola satisfies two such conditions. It is the locus of all points that have equal distance from one point and one straight line. The parabola may be called a compromise between two structural demands. Either structural condition yields to the other.<sup>3</sup> Is there any possible connection between these geometrical characteristics of the parabola and the particular configuration of physical forces to which we attribute gentleness? One may point to the kind of physical process that produces parabolic patterns. In ballistics, for instance, the parabolic curve of a trajectory is the result of a 'compromise' between the direction of the original impulse and the gravitational attraction. The two forces 'yield' to each other.<sup>4</sup>

At this point the description must shift from the observed person A to the observer B. B's eyes receive an image (level VI) of the gesture performed by A's arm and hand. Why should this

<sup>3</sup> One can express this also in terms of projective geometry by saying that the parabola as a conic section is intermediate between the horizontal section, namely the circle, and the vertical section, the straight-edged triangle.

<sup>4</sup> One of the principles on which the analysis of handwritings is based indicates that the script pattern reflects dynamic features of the writer's motor behavior, which in turn is produced by a characteristic configuration of muscular forces. The same isomorphism of muscular behavior and resulting visible trace has found applications in the technique of drawing. Langfeld (15, p. 129) quotes Bowie (5, pp. 35 and 77-79) concerning the principle of 'living movement' (*Sei Do*) in Japanese painting: "A distinguishing feature in Japanese painting is the strength of the brush stroke, technically called *fude no chikara* or *fude no ikioi*. When representing an object suggesting strength, such, for instance, as rocky cliff, the beak or talons of a bird, the tiger's claws, or the limbs and branches of a tree, the moment the brush is applied the sentiment of strength must be invoked and felt throughout the artist's system and imparted through his arm and hand to the brush, and so transmitted into the object painted."

image produce in B the impression that he is observing a gentle gesture? It may be true that the geometrical pattern of the gesture as well as the configuration of muscular forces which has created this pattern can both be characterized structurally as containing compromise, flexibility, yielding. But this fact in itself is not sufficient to explain the direct experience which B is said to receive by his perceptual observation. It becomes clear at this point that the Gestalt theory of expression is faced not only with the problem of showing how psychical processes can be inferred from bodily behavior, but that the primary task consists in making plausible the fact that the perception of shape, movement, etc. may convey to the observer the direct experience of an expression which is structurally similar to the organization of the observed stimulus pattern.

A's gesture is projected on the retinae of B's eyes<sup>5</sup> and, by way of the retinal images, on the visual cortex of B's cerebrum (level VII). Correspondingly, B perceives A's gesture (level VIII). Is there a possible similarity of the geometrical structure of the stimulus configuration and the structure of the expression which it conveys to the observer? We may go back to our mathematical analysis of the circle and the parabola. Simple experiments confirm what artists know from experience, namely that a circular curve looks 'harder,' less flex-

<sup>5</sup> At this stage a number of factors may interfere with the adequate projection of decisive characteristics of body A on the receptor organ of B. In our specific example it will depend, for instance, on the angle of projection, whether or not the perspective retinal image will preserve the essential structural features of the parabolic movement or transform it into a stimulus trace of unclear or clearly different structure. (In photographs and motion pictures such factors influence the kind of expression obtained from the reproduction of physical objects.) Similar factors will influence the veracity of other perceptual qualities which carry expression.

ible, than a parabolic one. In comparison with the circle the parabola looks more gentle. One could try to explain this finding by assuming that the observer knows, through past experience, the geometrical characteristics of such patterns or the nature of the physical forces which frequently produce them. This would take us back to the associationist theory. Along Gestalt lines another explanation suggests itself.

The projection of the perceptual stimulus on the visual cortex can be assumed to create a configuration of electro-chemical forces in the cerebral field. The well-known Gestalt experiments in perception suggest that retinal stimulations are subjected to organizational processes when they reach the cortical level. As a result of these processes the elements of visual patterns are perceived as being grouped according to Wertheimer's rules. Furthermore, any visual pattern appears as an organized whole, in which some predominant elements determine the overall shape and the directions of the main axes, while others have subordinate functions. For the same reasons, modifications of objective shape and size are perceived under certain conditions.

It will be observed that all these experimental findings focus upon the effects of the strains and stresses which organize the cortical field. Is there any reason to assume that only the effects of these dynamic processes, namely the groupings, the hierarchies of structural functions, and the modifications of shape and size, are reflected in perceptual experience? Why should not the strains and stresses of the cortical forces themselves also have their psychological counterpart? It seems plausible that they represent the physiological equivalent of what is experienced as expression.

Such a theory would make expression an integral part of the elementary processes of perception. *Expression, then,*

*could be defined as the psychological counterpart of the dynamic processes which result in the organization of perceptual stimuli.* While concrete verification is obviously far away, the basic assumption has gained in concreteness since Köhler and Wallach (14) have explained phenomena of perceptual size, shape and location through the action of electro-chemical forces. The future will show whether the theory can be extended to covering the phenomena of expression.

It is possible now to return to the question of how the perception of shape, movement, etc. may convey to an observer the direct experience of an expression which is structurally similar to the organization of the observed stimulus pattern. We referred previously to the constellations of physical forces which will induce an object to pursue a parabolic path. The physicist may be able to tell whether the example from ballistics is invertible. Will a parabolic pattern, such as the one projected on the cortical field, under certain conditions set off a configuration of forces which contains the structural factors of 'compromise' or 'yielding'? If so, isomorphism of the cortical forces and those described as levels I-V could be established.

This brings the description of isomorphic levels to an end. If the presentation is correct, the Gestalt-theoretical thesis would imply that an observer will adequately gauge another person's state of mind by inspection of that person's bodily appearance if the psychical situation of the observed person and the perceptual experience of the observer are structurally similar by means of a number of intermediate isomorphic levels.

*Expression as a perceptual quality.* The definition which was given above suggests that expression is an integral part of the elementary perceptual proc-

ess. This should not come as a surprise. Perception is a mere instrument for the registration of color, shape, sound, etc. only as long as it is considered in isolation from the organism, of which it is a part. In its proper biological context, perception appears as the means by which the organism obtains information about the friendly, hostile, or otherwise relevant environmental forces to which it must react. These forces reveal themselves most directly by what is described here as expression.

There is psychological evidence to bear out this contention. In fact, the observations on primitives and children cited by Werner (21, pp. 67-82) and Köhler (13) indicate that 'physiognomic qualities,' as Werner calls them, are even more directly perceived than the 'geometric-technical' qualities of size, shape or movement. Expression seems to be the primary content of perception. To register a fire as merely a set of hues and shapes in motion rather than to experience primarily the exciting violence of the flames presupposes a very specific, rare and artificial attitude. Even though the practical importance of, and hence the alertness to, expression has decreased in our culture, it cannot be maintained that a basic change has taken place in this respect. Darwin (7, pp. 359-360) noted that people sometimes observe and describe facial expression without being able to indicate the features of form, size, direction, etc. which carry it. In experimental work, one notices that even with the object directly in front of their eyes, subjects find it a hard and uncomfortable task to take note of the formal pattern. They constantly fall back upon the expressive characteristics, which they describe freely and naturally. Everyday experience shows that people may clearly recall the expression of persons or objects without being able to indicate color or shape. Asch observes: "Long before one

has realized that the color of the scene has changed, one may feel that the character of the scene has undergone change" (3, p. 85). Finally, there is the fact that the artist's, writer's, musician's approach to their subject is principally guided by expression.<sup>6</sup>

*Generalized theory.* Thus far, the phenomenon of expression has been discussed essentially in its best known aspect, namely, as a physical manifestation of psychical processes. However, some of the foregoing considerations implied that expression is a more universal phenomenon. Expression does not only exist when there is a mind 'behind' it, a puppeteer that pulls the strings. Expression is not limited to living organisms, which possess consciousness. A flame, a tumbling leaf, the wailing of a siren, a willow tree, a steep rock, a Louis XV chair, the cracks in a wall, the warmth of a glazed teapot, a hedgehog's thorny back, the colors of a sunset, a flowing fountain, lightning and thunder, the jerky movements of a bent piece of wire—they all convey expression through the various senses. The importance of this fact has been concealed by the popular hypothesis that in such cases human expression is merely transferred to objects. If, however, expression is an inherent characteristic of perceptual factors, it becomes unlikely that non-human expression should be nothing but an anthropomorphism, a 'pathetic fallacy.' Rather will human expression have to be considered a special case of a more general phenomenon. The comparison of an object's expression with a human state of mind is a secondary process (*cf.* p. 165). A weeping willow does not look sad because it looks like a sad person. It is more adequate to state that since the shape, direction and flexibility of willow branches convey the expression of passive hanging, a comparison with

<sup>6</sup> This has led to the erroneous notion that all perception of expression is aesthetic.

the structurally similar psycho-physical pattern of sadness in humans may impose itself secondarily.

Expression is sometimes described as 'perceiving with imagination.' In doing so Gottschalk (9) explains that "something is perceived as if it were actually present in the object of perception, although literally it is only suggested and not actually there. Music is not literally sad or gay or gentle; only sentient creatures or creatures with feeling, such as human beings, could be that." If our language possessed more words which could refer to kinds of expression as such, instead of naming them after emotional states in which they find an important application, it would become apparent that the phenomenon in question is "actually present in the object of perception" and not merely associated with it by imagination.

Even with regard to human behavior, the connection of expression with a corresponding state of mind is not as compelling and indispensable as is sometimes taken for granted. Köhler (12, pp. 260-264) has pointed out that people normally deal with and react to the expressive physical behavior in itself rather than being conscious of the psychical experiences reflected by such behavior. We perceive the slow, listless, 'droopy' movements of one person as against the brisk, straight, vigorous movements of another, but do not necessarily go beyond the meaning of such appearance by thinking explicitly of the psychical weariness or alertness behind it. Weariness and alertness are already contained in the physical behavior itself; they are not distinguished in any essential way from the weariness of slowly floating tar or the energetic ringing of the telephone bell.

This broader conception has practical consequences. It suggests, for instance, that the phenomenon of expression does not belong primarily under the heading

of the emotions or personality, where it is commonly treated. It is true that the great contributions which the study of expression has in store for these fields of psychology are thus far almost untapped. However, the experience of the last decades shows that little progress is made unless the nature of expression itself is clarified first.<sup>7</sup>

*Secondary effects.* Strictly speaking, the phenomenon of expression is limited to the levels V-VIII of Table 1. That is, the term 'expression,' as used in this paper, refers to an experience which takes place when a sensory stimulus affects the visual cortex of an observer's brain. The processes which may have given rise to the stimulus as well as those which the cortical stimulation provokes in other brain centers are supplementary.

Once perceptual stimulation has taken place, a number of secondary happenings may follow. (1) The observer B may deduce from the expression of B's bodily behavior that particular psychical processes are going on in A's mind; that is, through the perception of level V the observer gains knowledge about level I. The observation of a gentle gesture leads to the conclusion: B is in a gentle mood. This conclusion may be based on an isomorphic similarity between the observed behavior and a state of mind known or imaginable to the observer. In other cases, the conclusion may rely on past experience. Yawning, for instance, conveys the direct expression of sudden expansion; but the connection between yawning and fatigue or boredom is discovered by learning. The same seems to be true for the spasmodic outbursts of sound which we call laughter and which

<sup>7</sup> Once this is done, it will be possible and necessary to approach the further problem of the influences which the total personality exerts upon the observation of expression. To Vincent van Gogh, cypress trees conveyed an expression which they do not have for many other people. Cf. Koffka (10, p. 600).

in themselves are so far from suggesting mirth that they remain permanently incomprehensible to the chimpanzee, who otherwise "at once correctly interprets the slightest change of human expression, whether menacing or friendly" (11, p. 307). It is important to realize that an expression may be correctly perceived and described, yet the inferences derived from it may be wrong. If, in an experiment, 80 per cent of the observers agree on an 'erroneous' attribution, it is not sufficient to dismiss the result as an instance of failure. The high amount of agreement represents a psychological fact in its own right. The reliability of the observers' responses to a perceptual stimulus is a problem quite different from the validity of such responses, *i.e.*, the question whether the observers' diagnosis is 'true.'

(2) The observed expression may bring about the corresponding state of mind in B. In perceiving A's gentle behavior, the observer himself may experience a feeling of tenderness. (Lipps speaks of 'sympathetic empathy' as distinguished from 'simple empathy' 16, p. 417). (3) The observed expression may provoke the corresponding kinesthetic experience, *e.g.*, a feeling of relaxed softness. The effects described under (2) and (3) may be instances of a kind of 'resonance' based on isomorphism. Just as a sound calls forth a vibration of similar frequency in a string, various levels of psychological experience, such as the visual, the kinesthetic, the emotional seem to elicit in each other sensations of similar structure. (4) The perceived expression may remind B of other observations in which a similar expression played a role. Thus past experience is considered here not as the basis for the apperception of expression; instead, the direct observation of expression becomes the basis for comparison with similar observations in the past.

*The role of past experience.* While

there is no evidence to support the hypothesis that the central phenomenon of expression is based on learning, it is worth noting that in most cases the interpretation of the perceived expression is influenced by what is known about the person or object in question and about the context in which it appears. Mere inspection will produce little more than overall impressions of the forces at work, strong and clear-cut as such an experience may be. Increasing knowledge will lead to more differentiated interpretations, which will take the particular context into account. (As an example, one may think of the expression conveyed by the behavior of an animal whose habits one does not know and the changes that occur with closer acquaintance.) Knowledge does not interfere with expression itself, it merely modifies its interpretation, except for cases in which knowledge changes the appearance of the carrier of expression, that is, the perceptual pattern itself. For instance, a line-figure may change its perceptual structure and therefore its expression if it is suddenly seen as a human figure. A lifted eyebrow is seen as tense because it is perceived as a deviation from a known normal position. The expression of Mongolian eyes or Negro lips is influenced, for a white observer, by the fact that he conceives them as deviations from the normal face of his own race.

In Gestalt terms, past experience, knowledge, learning, memory are considered as factors of the temporal context in which a given phenomenon appears. Like the spatial context, on which Gestaltists have concentrated their attention during the early development of the theory, the temporal context influences the way a phenomenon is perceived. An object looks big or small depending on whether it is seen, spatially, in the company of smaller or larger objects. The same is true for the temporal context. The buildings of a

middle-sized town look tall to a farmer, small to a New Yorker, and correspondingly their expression differs for the two observers. Mozart's music may appear serene and cheerful to a modern listener, who perceives it in the temporal context of twentieth-century music, whereas it conveyed the expression of violent passion and desperate suffering to his contemporaries against the background of the music they knew. Such examples do not demonstrate that there is no intrinsic connection between perceptual patterns and the expression they convey but simply that experiences must not be evaluated in isolation from their spatial and temporal whole-context.

Knowledge often merges with directly perceived expression into a more complex experience. When we observe the gentle curve of a coachman's whip while being aware at the same time of the aggressive use of the object, the resulting experience clearly contains an element of contradiction. Such contradictions are exploited by artists; compare, in motion pictures, the uncanny effect of the murderer who moves softly and speaks with a velvety voice.

Finally, the perceptual experience of expression can be influenced by the kind of training which in artistic and musical instruction is known as making students 'see' and 'hear.' By opening people's eyes and ears to what is directly perceptible, they can be made to scan the given sensory pattern more adequately and thus to receive a fuller experience of its expression. A neglected or misled capacity for responding perceptually can be revived or corrected.

*The role of kinesthesia.* Frequently people feel that another person, whom they are observing, behaves physically the way they themselves have behaved before. They get this impression even though at that time they probably did not watch themselves in the mirror. It may be that they compare their own

state of mind as they remember it from the former occasion with the expression conveyed by the bodily behavior of the other person and/or with the state of mind reflected in that behavior. Probably the kinesthetic perception of one's own muscular behavior plays an important part in such situations. If muscular behavior and kinesthetic experience are isomorphic, it becomes explainable why at times one is so keenly aware of one's own facial expression, posture, gestures. One may feel, for instance: Right now, I look just like my father! The most convincing example is furnished by actors and dancers, whose bodily performance is created essentially through kinesthetic control. And yet their gestures are understandable to the audience visually. This suggests that there is a valid correspondence between bodily behavior and the related kinesthetic perception. The problem of what enables an infant to imitate an observer who smiles or shows the tip of his tongue belongs in the same category. Of particular interest is the fact that the blind express their feelings—even though imperfectly—in spite of their inability to observe expression in others visually. The blind also understand certain gestures on the basis of their own kinesthetic experiences.

"The blind man, like the person who sees, is aware of the gestures he makes when under the influence of various emotions. He shrugs his shoulders and raises his arms to express his disdain and amazement. The same gestures recognized by him in a statue will evoke within him the same sentiments" (20, p. 320).

Isomorphism would seem to account also for the fact that it often suffices to assume a particular posture (levels III and IV) in order to enter into a corresponding state of mind (level I). Bending the head and folding the hands is more than an accidentally chosen pos-

ture of praying, which derives its meaning merely from tradition. The kinesthetic sensation which accompanies this posture is structurally akin to the psychical attitude called devotion. 'Bow-ing' to a superior power's will is a mental condition so directly related to the corresponding bodily gesture that its common linguistic description uses the physical to describe the psychological. Rituals not only express what people feel but also help them to feel the way the situation requires. By straightening our backbones we produce a muscular sensation which is akin to the attitude of pride, and thus introduce into our state of mind a noticeable element of bold self-sufficiency.<sup>8</sup>

Even the 'practical' motor activities are accompanied more or less strongly by structurally corresponding states of mind. For instance, hitting or breaking things normally seems to evoke the emotional overtone of attack. To assert merely that this is so because people are aggressive would be an evasion of the problem. But if the dynamic character of the kinesthetic sensation which accompanies hitting and breaking resembles the emotional dynamics of attack, then the one may be expected to evoke the other—by 'resonance' (*cf. p. 166*). (This kinship makes it possible for aggressiveness, wherever it exists, to express itself through such motor acts.) Probably this parallelism holds true for all motor activity. Muscular behavior such as grasping, yielding, lifting, straightening, smoothing, loosening, bending, running, stopping seems to produce mental resonance effects constantly. (In consequence, language uses all of them metaphorically to describe states

<sup>8</sup> James's theory of emotion is based on a sound psychological observation. It fails where it identifies the kinesthetic sensation with the total emotional experience instead of describing it as a component which reinforces and sometimes provokes emotion because of the structural similarity of the two.

of mind.) The psychosomatic phenomena of pathological 'organ-speech' ("I cannot stomach this!") may be considered the most dramatic examples of a universal interdependence. The range and the importance of the phenomenon are not acknowledged as long as one studies expression only in motor activities that are not, or not any more, serviceable. It seems safe to assert that all motor acts are expressive, even though in different degrees, and that they all carry the experience of corresponding higher mental processes, if ever so faintly. Therefore, it is inadequate to describe expressive movements as mere atavisms, the way Darwin did. They are physical acts which take place because of their inner correspondence with the state of mind of the person who performs them. To use one of Darwin's examples: a person who coughs in embarrassment is not simply the victim of a meaningless association between a state of mind and a physical reaction, which was or can be serviceable under similar circumstances. Rather does he produce a reaction which he experiences to be meaningfully related to his state of mind. The bodily accompaniment completes the mental reaction. Together they form an act of total psycho-physical behavior. The human organism always functions as a whole, physically and psychically.

This view permits an application to the theory of art. It highlights the intimate connection of artistic and 'practical' behavior. The dancer, for instance, does not have to endow movements with a symbolic meaning for artistic purposes, but uses in an artistically organized way the unity of psychical and physical reaction which is characteristic for human functioning in general.

In a broader sense, it is the direct expressiveness of all perceptual qualities which allows the artist to convey

the effects of the most universal and abstract psycho-physical forces through the presentation of individual, concrete objects and happenings. While painting a pine tree, he can rely on the expression of towering and spreading which this tree conveys whenever it is seen by a human eye, and thus can span in his work the whole range of existence, from its most general principles to the tangible manifestations of these principles in individual objects.

*An illustration.* It has been pointed out in the beginning that experimenters have been concerned mostly with the question whether and to what extent observers can judge a person's state of mind from his physical appearance. In consequence, the psychological literature contains few analyses of perceptual patterns with regard to the expression they convey. As an example of the kind of material which is badly needed in this field, Efron's study on the gestures of two ethnical groups (8) may be cited. He describes the behavior of Eastern Jews and Southern Italians in New York City by analyzing the range,

speed, plane, coordination, and shape of their movements. A comparison of these findings with the mentalities of the two groups would probably produce excellent illustrations of what is meant by the structural similarity of psychical and physical behavior. Among the experimental investigations, Lundholm's early study (17) may be mentioned. He asked eight laymen in art to draw lines, each of which was to express the affective tone of an adjective given verbally. It was found, for instance, that only straight lines, broken by angles, were used to represent such adjectives as exciting, furious, hard, powerful, while only curves were used for sad, quiet, lazy, merry. Upward direction of lines expressed strength, energy, force; downward direction, weakness, lack of energy, relaxation, depression, etc. Recently Willmann (22) had thirty-two musicians compose short themes, meant to illustrate four abstract designs. Some agreement among the composers was found concerning the tempo, meter, melodic line, and amount of consonance, chosen to render the characteristics of

TABLE 2  
ANALYSIS OF DANCE MOVEMENTS IMPROVISED BY FIVE SUBJECTS

	<i>Sadness:</i>	<i>Strength:</i>	<i>Night:</i>
<i>Speed:</i>	5: slow	2: slow 1: very fast 1: medium 1: decrescendo	5: slow
<i>Range:</i>	5: small, enclosed	5: large, sweeping	3: small 2: large
<i>Shape:</i>	3: round 2: angular	5: very straight	5: round
<i>Tension:</i>	4: little tension 1: inconsistent	5: much tension	4: little tension 1: decrescendo
<i>Direction:</i>	5: indefinite, changing, wavering	5: precise, sharp, mostly forward	3: indefinite, changing 2: mostly downward
<i>Center:</i>	5: passive, pulled downward	5: active, centered in body	3: passive 2: from active to passive

the drawings. Subsequently the designs and compositions were used for matching experiments.

Because of the scarcity of pertinent material, it may be permissible to mention here an experiment which is too limited in the number of cases and too subjective in its method of recording and evaluating the data to afford a proof of the thesis we are discussing. It is presented merely as an example of the kind of research which promises fruitful results.<sup>9</sup> Five members of the student dance group of Sarah Lawrence College were asked individually to give improvisations of the following three subjects: sadness, strength, night. Rough descriptions of the dance patterns which resulted were jotted down by the experimenter and later classified according to a number of categories. Table 2 presents the findings in an abbreviated form. The numerical agreement is high but obviously carries little weight. As a point of method, it may only be mentioned that instances of disagreement cannot be taken simply to indicate that there was no reliable correspondence between task and performance. Sometimes, the task allows more than one valid interpretation. For instance, 'strength' expresses itself equally well in fast and in slow movement. 'Night' is less directly related to one particular dynamic pattern than sadness or strength.

Most tempting is the comparison between the movement patterns and the corresponding psychical processes. Such comparison cannot be carried through with exactness at this time mainly because psychology has not yet provided a method of describing the dynamics of states of mind in a way which would be more exact scientifically than the descriptions offered by novelists or every-

<sup>9</sup> The data were collected and tabulated by Miss Jane Binney, a student at Sarah Lawrence College.

day language. Nevertheless, it can be seen from our example that the dynamic patterns of expressive behavior permit relatively concrete and exact descriptions in terms of speed, range, shape, etc. Even the crudely simplified characterizations given in the table seem to suggest that the motor traits through which the dancers interpreted sadness reflect the slow, languishing pace of the psychological processes, the indefiniteness of aim, the withdrawal from the environment, the passivity—all of which distinguish sadness psychologically. The fact that expressive behavior is so much more readily accessible to concrete scientific description than the corresponding psychical processes deserves attention. It suggests that in the future the study of behavior may well become the method of choice, when psychologists undertake the task of reducing emotions and other psychical processes to configurations of basic forces. Already the analysis of handwriting has led to a number of categories (pressure, size, direction, proportion, etc.) which invite a search for the corresponding psychological concepts.

Our example will also show why it is fruitless to dismiss the phenomena of expression as 'mere stereotypes.' If it can be demonstrated that the dynamics of psychical and physical processes are structurally interrelated and that this interrelation is perceptually evident, the question of whether and to what extent the performance and its interpretations are based on social conventions loses importance.

#### BIBLIOGRAPHY

1. ALLPORT, G. W., & VERNON, P. E. *Studies in expressive movement*. New York: Macmillan, 1933.
2. ARNHEIM, R. Experimentell-psychologische Untersuchungen zum Ausdrucksproblem. *Psychol. Forsch.*, 1928, 11, 2-132.

3. ASCH, S. E. Max Wertheimer's contribution to modern psychology. *Social Research*, 1946, 13, 81-102.
4. BERKELEY, G. *An essay toward a new theory of vision*. New York: Dutton, 1934.
5. BOWIE, H. P. *On the laws of Japanese painting*. San Francisco: Elder, 1911.
6. BUEHLER, C. The social behavior of children. In C. Murchison (Ed.), *A handbook of child psychology*. Worcester: Clark, 1933. Pp. 374-416.
7. DARWIN, C. *The expression of the emotions in man and animals*. New York: Appleton, 1896.
8. EFRON, D. *Gesture and environment*. New York: King's Crown, 1941.
9. GOTTSCHALK, D. W. *Art and the social order*. Chicago: Univ. Press, 1947.
10. KOFFKA, K. *Principles of Gestalt psychology*. New York: Harcourt Brace, 1935.
11. KÖHLER, W. *The mentality of apes*. New York: Harcourt Brace, 1925.
12. —. *Gestalt psychology*. New York: Liveright, 1929.
13. —. Psychological remarks on some questions of anthropology. *Amer. J. Psychol.*, 1937, 50, 271-288.
14. —, & WALLACH, H. Figural after-effects. An investigation of visual processes. *Proc. Amer. phil. Soc.*, 1944, 88, 269-357.
15. LANGFELD, H. S. *The aesthetic attitude*. New York: Harcourt Brace, 1920.
16. LIPPS, T. Aesthetische Einfühlung. *Z. Psychol.*, 1900, 22, 415-450.
17. LUNDHOLM, H. The affective tone of lines. *PSYCHOL. REV.*, 1921, 28, 43-60.
18. RAPAPORT, D. The Szondi test. *Bull. Menninger Clin.*, 1941, 5, 33-39.
19. SCHACHTEL, E. G. On color and affect. *Psychiatry*, 1943, 6, 393-409.
20. VILLEY, P. *The world of the blind*. London: Duckworth, 1930.
21. WERNER, H. *Comparative psychology of mental development*. New York: Harper, 1940.
22. WILLMANN, R. R. An experimental investigation of the creative process in music. *Psychol. Monogr.*, 1944, 57, No. 261.
23. WOLFF, W. *The expression of personality*. New York: Harper, 1943.
24. WOODWORTH, R. S. *Experimental psychology*. New York: Holt, 1939.

[MS. received October 20, 1948]

## 'SUPERSTITIOUS' BEHAVIOR IN ANIMALS

BY W. N. KELLOGG

*Indiana University*

In an article entitled, "Superstition in the pigeon," Skinner (1) has recently pointed out that 'random' or irrelevant movements, if reinforced by a food stimulus, will become conditioned to that stimulus. The situation was one described by Skinner as operant conditioning. Small quantities of food were presented to the bird at regular time intervals without reference at all to the bird's behavior at the moment of presentation. If, by coincidence, the pigeon happened to be stamping its feet or tossing its head when the food was delivered, the stamping (or the tossing) behavior was reinforced by the food, and tended to recur on subsequent trials, prior to the presentation of the food.

It is the purpose of the present remarks to show that reactions of a similar nature can occur in animals other than the pigeon, as well as in situations which are not necessarily instances of operant conditioning. And also that the interpretations which are put upon them in the current state of our psychological thinking are likely to be conflicting.

Probably the earliest recorded case of persisting irrelevant responses of this sort in a *bona fide* experiment on animal learning is the behavior of Julius, the orang-utan, as described by Yerkes in 1916 (4). The problem was to enter the second-from-the-extreme-right of a row of boxes or stalls arranged before the animal in the multiple-choice experiment. The following statements from Yerkes describe 'the superstition' which developed (4, p. 82).

"A curiously interesting bit of behavior appeared for the first time on June 29. Julius had gone to the first box at the right end of the group, and instead of entering,

he had wheeled around toward his right, and turning a complete circle, faced the right [correct] box, which he promptly entered. Subsequently, the tendency developed and the method was used with increasing frequency . . .

"This odd bit of behavior proved peculiarly interesting and significant in that the tendency to turn became dissociated from the position (in front of the first box at the right end of the group) in connection with which it originally developed. After a few days, Julius would enter the reaction-chamber and instead of proceeding directly to the right end of the group, would stop suddenly wherever he happened to be, turn toward his right in a complete circle, and hasten into the box nearest to him which, as often as not, proved to be the wrong one."

One is reminded, by this description, of the losing player in a game of cards, who walks around his chair for luck, on the half-humorous assumption that this procedure will somehow improve the cards which he receives in future hands and so help him to win. In the case of Julius it is clear, however, that the persistence of the turning depended upon the early reinforcements it had received from his obtaining food in the correct box following its occurrence.

Similar instances of the reinforcing of irrelevant responses have been observed in our own laboratory for some years past in the flexion conditioning of dogs. The conditioned stimulus was a buzz of 1000 cycles which lasted on any trial for a period of 2.0 sec. The unconditioned or reinforcing stimulus was a momentary electric shock delivered to one of the feet upon the termination of the buzz. The procedure was one of instrumental avoidance conditioning in

which the animal could escape the shock on any trial by making a conditioned foot-lift to the buzz signal. By graphically recording the respiration as well as the responses of all four of the feet during the flexion-conditioning of only one of them, the pattern of activity which developed—so far as the limbs and the breathing were concerned—could be definitely measured from kymographic tracings.

It is typical for many dogs in this situation to go into a kind of struggle-spasm in response to the buzz during the early conditioning trials. The conditioned behavior at the start is likely to be generalized, therefore, and involves miscellaneous muscle tensions and the multiple flexing of all four of the limbs. This, of necessity, includes the flexing of the proper member, that is to say, it includes the response which from the point of view of the experimenter is 'the reaction to be conditioned.' As the training progresses much of the irrelevant activity is eliminated, but the process of elimination is often far from complete and sometimes reaches a kind of stable level which is maintained indefinitely.

If the right-rear foot is the foot which is given the reinforcing shock, an animal which has received 100 or so conditioning trials may evolve a characteristically individual pattern of responding which includes the other three feet as well. As soon as the buzz spurs him to action on any trial he may, for example, (a) raise the right-front foot and immediately return it to its original position, (b) raise the left-rear foot and replace it at once also, (c) raise and replace the left-front foot, and finally (d) lift and hold up the right-rear (or shocked) member during the time when the shock would have occurred had he failed to respond. Almost any other reaction, or combination of reactions, within the repertoire of a dog confined

in the conditioning stock, may come to be uniformly repeated as a preliminary to the flexing of the shocked limb itself.

The more readily observable of the 'superstitious' activities which have been noted in our laboratory during the flexion conditioning of dogs are listed below.

A. Graphically recorded 'superstitions.'

1. Movements of all four limbs.
2. Movements of three of the limbs.
3. Responses involving only two of the limbs.
4. Respiratory changes including those produced by barking or vocalizing.

B. 'Superstitions' not graphically recorded.

1. Head movements.
2. Chewing or biting of the conditioning stock.
3. Trunk and tail movements.
4. Postural adjustments.

Sample kymographic activity-records showing irrelevant conditioned response-patterns in all four legs, in three legs, and in two legs only, have, in fact, been published by Wolf and Kellogg (3) as far back as 1940. Such patterns of behavior seem to be developed 'accidentally' by the animal. If they occur during the buzz-period preceding the shock, they are thereby reinforced by the shock along with the basic conditioned-flexion response. Although characteristically different from animal to animal, they all contain as an essential and final component the response of the shocked member itself. The irrelevant movements in this way become a permanent part of the sequence of activity elicited by the buzz-stimulus. They frequently show no signs of extinction unless the basic component of the pattern is itself extinguished.

In the same way we can say, I think, that the 'superstitious' reactions of Skinner's pigeons were patterns of behavior ending with pecking and eating which were the basic or final responses of the sequence. The conditioned stimulus for the pattern as a whole was the regular time interval (in most cases 15 sec.) to which the bird began to react on any trial with its preliminary irrelevant movements. The unconditioned stimulus was the presentation of the food. The situation can thus, if one chooses, be regarded as an example of traditional or respondent conditioning, rather than an example of operant conditioning.

By calling behavior of this sort 'superstitious' one implicitly defines the particular pattern of activity which occurs before the final response as 'the superstition.' This use of the term also implies—by a kind of whimsical analogy to human experience—that the organism 'believes' it is 'causing' the food to appear as a result of its reactions. A human observer, noting the repeated occurrence of these complicated maneuvers upon each trial, may tend to assume that they are mystically indulged in by the animal for the purpose of producing the food in the case of the pigeons, or of removing the shock in the case of the dogs. They are likely, therefore, to be thought of as the means by which the animal seeks to solve the experimental problem. By adding the concept of causality to these behavior-patterns in this manner, one makes of them superstitions in the literal sense, since the organism is now conceived of as ascribing unreal or unnatural causes to a straightforward series of events. Rather than being ordinary instances of the association of movements in series, they are looked upon as cognitive or purposive acts.

Just how many, if any, of these so-called superstitious activities would be

interpreted as having causal properties from the point of view of the animal, must depend upon the theoretical predisposition of the individual observer. The background for this position as well as for that of the opposing view, can be traced to the old conflict between mentalistic and objective psychology—to the long history of purposive versus incidental or mechanical explanations. The mentalistic or cognitive interpretation, since it is applied in this instance to animals, is anthropomorphic. The associative or behavioristic interpretation is, by contrast, mechanomorphic (2).

#### SUMMARY

The points brought out in the foregoing discussion may be summarized as follows:

1. Animals other than the pigeon will behave 'superstititiously.'
2. 'Superstitious' behavior will appear in dogs, for example, in the buzz-shock conditioning situation. It was found by Yerkes in the multiple-choice experiment with an orang-utan. It is highly probable that it will also occur in many other learning situations (and with many other animals) in which irrelevant behavior is either positively or negatively reinforced.
3. Whether such behavior in animals is regarded as literally superstitious in the sense that it is indulged in as a causal or purposive maneuver to produce the food or eliminate the shock, is a matter which depends upon the interpretation read into it by one observer or another. The causal explanation must be regarded as mentalistic and anthropomorphic. An alternate and less complicated interpretation would be that these activities—whether relevant or irrelevant to the solution of the problem—are ordinary instances of the association of a sequence of movements with a reinforcing stimulus.

## REFERENCES

1. SKINNER, B. F. 'Superstition' in the pigeon. *J. exp. Psychol.*, 1948, **38**, 168-172.
2. WATERS, R. H. Mechanomorphism: a new term for an old mode of thought. *Psychol. Rev.*, 1948, **55**, 139-142.
3. WOLF, I. S., & KELLOGG, W. N. Changes in general behavior during flexion condi-
- tioning and their importance for the learning process. *Amer. J. Psychol.*, 1940, **43**, 384-396.
4. YERKES, R. S. The mental life of monkeys and apes: a study of ideational behavior. *Behav. Monogr.*, 1916, **3**, No. 12, 145 p.

[MS. received October 20, 1948]



# REPORT TO LIBRARIANS

*concerning*

## Available Back Issues

*of the*

## Psychological Review

Thirty-two issues of the Psychological Review became out-of-print in 1948. Libraries and institutions ordered 11 sets of "all available issues," and 32 issues had fewer than 11 copies.

The American Psychological Association does not plan to reprint missing issues. At the rate of sale established for 1948, 24 more issues will become out-of-print during 1949.

The list of available issues of the Psychological Review, as found on the back cover of this journal, agrees with the inventory of December 31, 1948.

# PSYCHOLOGICAL REVIEW

YEAR	VOLUME	AVAILABLE NUMBERS	PRICE PER NUMBER	PRICE PER VOLUME
1894	1	- 2	\$1.00	\$1.00
1895	2	- - 3	\$1.00	\$1.00
1896	3	- - -		
1897	4	1 - -	\$1.00	\$2.00
1898	5	- 2 3	\$1.00	\$3.00
1899	6	- - 4	\$1.00	\$1.00
1900	7	- - -	\$1.00	\$1.00
1901	8	1 - 2	\$1.00	\$2.00
1902	9	- 2 -	\$1.00	\$1.00
1903	10	1 - 2 3	\$1.00	\$3.00
1904	11	- 2 3 4	\$1.00	\$4.00
1905	12	1 - 2 3 4 5	\$1.00	\$5.00
1906	13	- 2 3 4 5 6	\$1.00	\$6.00
1907	14	1 - 2 3 4 5 6	\$1.00	\$6.00
1908	15	- - -		
1909	16	1 - -	\$1.00	\$5.00
1910	17	1 - 2 3 4 5 6	\$1.00	\$6.00
1911	18	1 - 2 3 4 5 6	\$1.00	\$5.50
1912	19	1 - 2 3 4 5 6	\$1.00	\$5.50
1913	20	1 - 2 3 4 5 6	\$1.00	\$5.50
1914	21	1 - 2 3 4 5 6	\$1.00	\$5.00
1915	22	1 - 2 3 4 5 6	\$1.00	\$5.50
1916	23	1 - 2 3 4 5 6	\$1.00	\$4.00
1917	24	- - -	\$1.00	\$1.00
1918	25	1 - 2 3 4 5 6	\$1.00	\$5.00
1919	26	1 - 2 3 4 5 6	\$1.00	\$5.50
1920	27	1 - 2 3 4 5 6	\$1.00	\$5.50
1921	28	1 - 2 3 4 5 6	\$1.00	\$4.00
1922	29	1 - 2 3 4 5 6	\$1.00	\$2.00
1923	30	1 - 2 3 4 5 6	\$1.00	\$5.50
1924	31	1 - 2 3 4 5 6	\$1.00	\$5.50
1925	32	- - -	\$1.00	\$3.00
1926	33	1 - 2 3 4 5 6	\$1.00	\$5.50
1927	34	1 - 2 3 4 5 6	\$1.00	\$5.50
1928	35	1 - 2 3 4 5 6	\$1.00	\$5.50
1929	36	1 - 2 3 4 5 6	\$1.00	\$5.50
1930	37	1 - 2 3 4 5 6	\$1.00	\$5.50
1931	38	1 - 2 3 4 5 6	\$1.00	\$5.50
1932	39	1 - 2 3 4 5 6	\$1.00	\$5.50
1933	40	1 - 2 3 4 5 6	\$1.00	\$5.50
1934	41	1 - 2 3 4 5 6	\$1.00	\$5.50
1935	42	1 - 2 3 4 5 6	\$1.00	\$5.50
1936	43	1 - 2 3 4 5 6	\$1.00	\$5.50
1937	44	1 - 2 3 4 5 6	\$1.00	\$5.50
1938	45	1 - 2 3 4 5 6	\$1.00	\$5.50
1939	46	1 - 2 3 4 5 6	\$1.00	\$5.50
1940	47	1 - 2 3 4 5 6	\$1.00	\$5.50
1941	48	1 - 2 3 4 5 6	\$1.00	\$5.50
1942	49	1 - 2 3 4 5 6	\$1.00	\$5.50
1943	50	1 - 2 3 4 5 6	\$1.00	\$5.50
1944	51	1 - 2 3 4 5 6	\$1.00	\$5.00
1945	52	1 - 2 3 4 5 6	\$1.00	\$5.50
1946	53	1 - 2 3 4 5 6	\$1.00	\$5.00
1947	54	1 - 2 3 4 5 6	\$1.00	\$5.50
1948	55	1 - 2 3 4 5 6	\$1.00	\$5.50
1949	56	By Subscription, \$5.50	\$1.00	

List price, Volumes 1 through 55

30% Discount

Net price, Volumes 1 through 55

\$241.50

72.45

\$169.05

Information about the Psychological Review: from 1894 to 1906, many numbers are out of print, as shown in the table. After 1906, twenty-three numbers are out of print. After 1925, two numbers are out of print.

The journal has been published with six numbers a year throughout its history.

Information about prices: the Psychological Review has the uniform price of \$5.50 per volume and \$1.00 per issue. For incomplete volumes, the price is \$1.00 for each available number. For foreign postage, \$2.25 per volume should be added. The American Psychological Association gives the following discounts on orders for any one journal:

10% on orders of \$50.00 and over

20% on orders of \$100.00 and over

30% on orders of \$150.00 and over

Current subscriptions and orders for back numbers should be addressed to  
AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.

1515 Massachusetts Avenue, N. W.

Washington 5, D. C.